

Nov 25 10:11:31 2003

us-09-724-685-107.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:05:35 ; Search time 33 Seconds  
(without alignments)  
2163.053 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949  
Sequence: 1 MWDGALPPTINSARYAGP.....SGVLAVPPRPVYMPHSPAAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1949	100.0	391	US-10-084-843-107	Sequence 107, App
2	1949	100.0	391	US-10-193-002-102	Sequence 102, App
3	1949	100.0	391	US-10-098-732A-14	Sequence 14, Appl
4	1949	100.0	596	US-09-287-849-26	Sequence 26, Appl
5	1949	100.0	596	US-10-359-460-26	Sequence 26, Appl
6	1949	100.0	596	US-10-098-732A-20	Sequence 20, Appl
7	1949	100.0	600	US-09-287-849-22	Sequence 22, Appl
8	1949	100.0	600	US-10-359-460-22	Sequence 22, Appl
9	1949	100.0	729	US-10-098-732A-18	Sequence 18, Appl
10	1949	100.0	930	US-10-098-732A-65	Sequence 65, Appl
11	1944	99.7	729	US-09-287-849-2	Sequence 2, Appl1
12	1944	99.7	729	US-10-359-460-2	Sequence 2, Appl1
13	1944	99.7	729	US-10-098-732A-16	Sequence 16, Appl
14	1652.5	84.8	396	US-10-084-843-111	Sequence 111, App
15	1652.5	84.8	396	US-10-193-002-106	Sequence 106, App

16	1486.5	76.3	359	12	US-10-084-843-109	Sequence 109, App
17	1486.5	76.3	359	12	US-10-193-002-104	Sequence 104, App
18	1187.5	60.9	358	9	US-09-287-849-8	Sequence 8, Appl1
19	1187.5	60.9	358	12	US-10-359-460-8	Sequence 8, Appl1
20	1187	60.9	263	12	US-10-084-843-91	Sequence 91, Appl
21	1187	60.9	263	12	US-10-193-002-92	Sequence 92, Appl
22	1187	60.9	263	12	US-10-098-732A-12	Sequence 12, Appl
23	766.5	39.3	400	9	US-09-073-009-126	Sequence 126, App
24	766.5	39.3	400	9	US-09-793-106-126	Sequence 126, App
25	731	37.5	421	16	US-10-080-170-146	Sequence 146, App
26	604	31.0	710	9	US-09-287-849-16	Sequence 16, Appl
27	604	31.0	710	12	US-10-359-460-16	Sequence 16, Appl
28	604	31.0	710	12	US-10-098-732A-49	Sequence 49, Appl
29	604	31.0	856	9	US-09-287-849-12	Sequence 12, Appl
30	603	30.9	856	12	US-10-359-460-12	Sequence 12, Appl
31	603	30.9	423	9	US-09-073-009-142	Sequence 142, App
32	603	30.9	423	9	US-09-793-106-142	Sequence 142, App
33	603	30.9	423	12	US-10-098-732A-31	Sequence 31, Appl
34	588	30.2	394	10	US-09-712-363-205	Sequence 205, App
35	439.5	22.6	655	10	US-09-712-363-207	Sequence 207, App
36	424.5	21.8	943	10	US-09-996-634-131	Sequence 131, App
37	424.5	21.8	943	11	US-09-997-182-131	Sequence 131, App
38	424.5	21.8	943	11	US-09-997-181-131	Sequence 131, App
39	424	21.8	141	9	US-09-073-009-15	Sequence 15, Appl
40	424	21.8	141	9	US-09-023-588-15	Sequence 15, Appl
41	424	21.8	141	9	US-09-793-106-15	Sequence 15, Appl
42	413	21.2	597	9	US-09-793-106-146	Sequence 146, App
43	364.5	16.3	408	16	US-10-080-170-57	Sequence 57, Appl
44	317	16.3	371	9	US-09-791-171-92	Sequence 92, Appl
45	317	16.3	371	12	US-09-804-980-92	Sequence 92, Appl

## ALIGNMENTS

RESULT 1  
US-10-084-843-107  
; Sequence 107, Application US/10084843  
; Publication No. US20030143243A1  
; GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
Skeiky, Yasir A.W.  
Dillon, Davin C.  
Campos-Neto, Antonio  
Houghton, Raymond  
Vedrick, Thomas S.  
Twardzik, Daniel J.  
Iodes, Michael J.  
Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA

ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/084,843  
FILING DATE: 25-Feb-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.

REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 107:  
US-10-084-843-107

Query Match 100.0%; Score 1949; DB 12; Length 391;  
Best Local Similarity 100.0%; Pred. No. 7.1e-146;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLWVAASPYYVWMSVTAGQAEITPAQVRAVAAAAYETAYGLTVPPIYAEKRAELMI 120  
DB 61 SSAGLWVAASPYYVWMSVTAGQAEITPAQVRAVAAAAYETAYGLTVPPIYAEKRAELMI 120  
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAODPAAAFGYAATATATATLLPFEAPEMTSAGG 180  
DB 121 LIATNLGQNTPAIAVNEAEYGEWMAODPAAAFGYAATATATATLLPFEAPEMTSAGG 180  
QY 181 LLEQAAVEEASDTPAANQLMNNVPQALQQLAOPGTGTPSSKLGKMTKTVSPHRSPISN 240  
DB 181 LLEQAAVEEASDTPAANQLMNNVPQALQQLAOPGTGTPSSKLGKMTKTVSPHRSPISN 240  
QY 241 MVSANNHMSMTNSGVSMNTTILSSMLKGFAPAAAQAQVTAONGVRAMSSLGSSSG 300  
DB 241 MVSANNHMSMTNSGVSMNTTILSSMLKGFAPAAAQAQVTAONGVRAMSSLGSSSG 300  
QY 301 LGGVVAANLGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
DB 301 LGGVVAANLGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
QY 361 GQMGARAGGSLGVLKRPVPPVPMHSPAG 391  
DB 361 GQMGARAGGSLGVLKRPVPPVPMHSPAG 391

## RESULT 2

US-10-193-002-102  
Sequence 102, Application US/10193002  
Publication No. US20030135026A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.

Skeiky, Yasir A.W.  
Dillon, David C.  
Campes-Neto, Antonia  
Houghton, Raymond  
Vedvick, Thomas S.  
Twardzik, Daniel R.  
Iodes, Michael J.  
Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
TUBERCULOSIS

NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESS: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/193,002  
FILING DATE: 10-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-10-193-002-102

Query Match 100.0%; Score 1949; DB 12; Length 391;  
Best Local Similarity 100.0%; Pred. No. 7.1e-146;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLWVAASPYYVWMSVTAGQAEITPAQVRAVAAAAYETAYGLTVPPIYAEKRAELMI 120  
DB 61 SSAGLWVAASPYYVWMSVTAGQAEITPAQVRAVAAAAYETAYGLTVPPIYAEKRAELMI 120  
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAODPAAAFGYAATATATATLLPFEAPEMTSAGG 180  
DB 121 LIATNLGQNTPAIAVNEAEYGEWMAODPAAAFGYAATATATATLLPFEAPEMTSAGG 180  
QY 181 LLEQAAVEEASDTPAANQLMNNVPQALQQLAOPGTGTPSSKLGKMTKTVSPHRSPISN 240  
DB 181 LLEQAAVEEASDTPAANQLMNNVPQALQQLAOPGTGTPSSKLGKMTKTVSPHRSPISN 240  
QY 241 MVSANNHMSMTNSGVSMNTTILSSMLKGFAPAAAQAQVTAONGVRAMSSLGSSSG 300  
DB 241 MVSANNHMSMTNSGVSMNTTILSSMLKGFAPAAAQAQVTAONGVRAMSSLGSSSG 300  
QY 301 LGGVVAANLGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
DB 301 LGGVVAANLGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
QY 361 GQMGARAGGSLGVLKRPVPPVPMHSPAG 391  
DB 361 GQMGARAGGSLGVLKRPVPPVPMHSPAG 391

## RESULT 3

US-10-098-732A-14  
Sequence 14, Application US/10098732A  
Publication No. US20030175294A1

GENERAL INFORMATION:  
APPLICANT: Skeiky, Yasir  
APPLICANT: Brannon, Mark  
APPLICANT: Gudexian, Jeffrey  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
FILE REFERENCE: 014058-012010US  
CURRENT APPLICATION NUMBER: US/10/098,732A  
CURRENT FILING DATE: 2003-04-29

;; PRIOR APPLICATION NUMBER: US 60/275,837  
;; PRIOR FILING DATE: 2001-03-13  
;; NUMBER OF SEQ ID NOS: 80  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 14  
;; LENGTH: 391  
;; TYPE: PR  
;; ORGANISM: Mycobacterium tuberculosis  
;; FEATURE:  
;; OTHER INFORMATION: MTB39 full length (TbH9FL)  
US-10-098-732A-14

Query Match 100.0%; Score 1949; DB 12; Length 391;  
Best Local Similarity 100.0%; Pred. No. 7,1e-146;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGWSIG 60  
DB 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGWSIG 60  
QY 61 SSAGLMVAAAAPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPVIAENRAELMI 120  
DB 61 SSAGLMVAAAAPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPVIAENRAELMI 120  
QY 121 LATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
DB 121 LATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
QY 181 LLEQAAAEEASDTAAANQLMNNVPOALQOQPTGTTSSKLGKMTVSPHRSPI 240  
DB 181 LLEQAAAEEASDTAAANQLMNNVPOALQOQPTGTTSSKLGKMTVSPHRSPI 240  
QY 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAVTAQNGVRAMSSIGSSIG 300  
DB 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAVTAQNGVRAMSSIGSSIG 300  
QY 301 LGGVAAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGPOMLGFLPV 360  
DB 301 LGGVAAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGPOMLGFLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391

RESULT 4  
US-09-287-849-26  
;; Sequence 26, Application US/09287849  
;; Patent No. US20020009459A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Reed, Steven G.  
;; APPLICANT: Skelky, Yasir A.W.  
;; APPLICANT: Dillon, Davin C.  
;; APPLICANT: Alderson, Mark  
;; APPLICANT: Campos-Neto, Antonio  
;; APPLICANT: Corixa Corporation  
;; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
;; FILE REFERENCE: 014058-009020US  
;; CURRENT APPLICATION NUMBER: US/09/287,849  
;; PRIOR FILING DATE: 1999-04-07  
;; PRIOR APPLICATION NUMBER: US 08/818,112  
;; PRIOR FILING DATE: 1997-03-13  
;; PRIOR APPLICATION NUMBER: US 08/942,578  
;; PRIOR FILING DATE: 1997-10-01  
;; PRIOR APPLICATION NUMBER: US 09/025,197  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 09/056,556  
;; PRIOR FILING DATE: 1998-04-07  
;; PRIOR APPLICATION NUMBER: US 09/223,040  
;; PRIOR FILING DATE: 1998-12-30  
;; NUMBER OF SEQ ID NOS: 46  
;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 26  
;; LENGTH: 596  
;; TYPE: PR  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: bi-fusion  
US-09-287-849-26

Query Match 100.0%; Score 1949; DB 9; Length 596;  
Best Local Similarity 100.0%; Pred. No. 1.2e-145;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGWSIG 60  
DB 9 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGWSIG 68  
QY 61 SSAGLMVAAAAPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPVIAENRAELMI 120  
DB 61 SSAGLMVAAAAPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPVIAENRAELMI 128  
QY 121 LATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
DB 121 LATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 188  
QY 181 LLEQAAAEEASDTAAANQLMNNVPOALQOQPTGTTSSKLGKMTVSPHRSPI 240  
DB 181 LLEQAAAEEASDTAAANQLMNNVPOALQOQPTGTTSSKLGKMTVSPHRSPI 248  
QY 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAVTAQNGVRAMSSIGSSIG 300  
DB 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAVTAQNGVRAMSSIGSSIG 308  
QY 301 LGGVAAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGPOMLGFLPV 360  
DB 301 LGGVAAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERGPOMLGFLPV 368  
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 399

RESULT 5  
US-10-359-460-26  
;; Sequence 26, Application US/10359460  
;; Publication No. US20030147911A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Reed, Steven G.  
;; APPLICANT: Skelky, Yasir A.W.  
;; APPLICANT: Dillon, Davin C.  
;; APPLICANT: Alderson, Mark  
;; APPLICANT: Campos-Neto, Antonio  
;; APPLICANT: Corixa Corporation  
;; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
;; FILE REFERENCE: 014058-009020US  
;; CURRENT APPLICATION NUMBER: US/10/359,460  
;; PRIOR FILING DATE: 2003-02-05  
;; PRIOR APPLICATION NUMBER: US/09/287,849  
;; PRIOR FILING DATE: 1999-04-07  
;; PRIOR APPLICATION NUMBER: US 08/818,112  
;; PRIOR FILING DATE: 1997-03-13  
;; PRIOR APPLICATION NUMBER: US 08/942,578  
;; PRIOR FILING DATE: 1997-10-01  
;; PRIOR APPLICATION NUMBER: US 09/025,197  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 09/056,556  
;; PRIOR FILING DATE: 1998-04-07  
;; PRIOR APPLICATION NUMBER: US 09/223,040  
;; PRIOR FILING DATE: 1998-12-30  
;; NUMBER OF SEQ ID NOS: 46  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 26  
;; LENGTH: 596

TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: bi-fusion  
US-10-359-460-26

Query Match 100.0%; Score 1949; DB 12; Length 596;  
Best Local Similarity 100.0%; Pred. No. 1.2e-145;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVSADLFSASAFOSVVMGLTVGSMIG 60  
DB 9 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVSADLFSASAFOSVVMGLTVGSMIG 68  
QY 61 SSAGLWVAASPVAWMSVTAGAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
DB 69 SSAGLWVAASPVAWMSVTAGAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMT 128  
QY 121 LIATNLGONTPALAIVNEAEYGEWMAODAAAMFGVAAATATATATLLPFEAEPMTSAGG 180  
DB 129 LIATNLGONTPALAIVNEAEYGEWMAODAAAMFGVAAATATATATLLPFEAEPMTSAGG 188  
QY 181 LLEQAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTIVSPHRSPISN 240  
DB 189 LLEQAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTIVSPHRSPISN 248  
QY 241 MVSANNNHSMNTNGSVMTNTLSMLKGFAPAAAQAVOTAAQNGVRAMSLGSSLGSSG 300  
DB 249 MVSANNNHSMNTNGSVMTNTLSMLKGFAPAAAQAVOTAAQNGVRAMSLGSSLGSSG 308  
QY 301 LGGGVAANLGRAASVGSLSVPOAAAQAVTPPAARALPLTSLTSAERPGQMLGGLPV 360  
DB 309 LGGGVAANLGRAASVGSLSVPOAAAQAVTPPAARALPLTSLTSAERPGQMLGGLPV 368  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
DB 369 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 399

RESULT 6  
US-10-098-732A-20  
Sequence 20, Application US/10098732A  
Publication No. US20030175294A1  
GENERAL INFORMATION:  
APPLICANT: Skeiky, Yasir  
APPLICANT: Brannon, Mark  
APPLICANT: Guderian, Jeffrey  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
FILE REFERENCE: 014058-012010US  
CURRENT APPLICATION NUMBER: US/10/098,732A  
CURRENT FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: US 60/275,837  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 20  
LENGTH: 596  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: bi-fusion  
OTHER INFORMATION: protein TbH9-Ra35 (designated MTB59F)  
US-10-098-732A-20

Query Match 100.0%; Score 1949; DB 12; Length 596;  
Best Local Similarity 100.0%; Pred. No. 1.2e-145;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVSADLFSASAFOSVVMGLTVGSMIG 60  
DB 9 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVSADLFSASAFOSVVMGLTVGSMIG 68

QY 61 SSAGLWVAASPVAWMSVTAGAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
DB 69 SSAGLWVAASPVAWMSVTAGAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMT 128  
QY 121 LIATNLGONTPALAIVNEAEYGEWMAODAAAMFGVAAATATATATLLPFEAEPMTSAGG 180  
DB 129 LIATNLGONTPALAIVNEAEYGEWMAODAAAMFGVAAATATATATLLPFEAEPMTSAGG 188  
QY 181 LLEQAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTIVSPHRSPISN 240  
DB 189 LLEQAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTIVSPHRSPISN 248  
QY 241 MVSANNNHSMNTNGSVMTNTLSMLKGFAPAAAQAVOTAAQNGVRAMSLGSSLGSSG 300  
DB 249 MVSANNNHSMNTNGSVMTNTLSMLKGFAPAAAQAVOTAAQNGVRAMSLGSSLGSSG 308  
QY 301 LGGGVAANLGRAASVGSLSVPOAAAQAVTPPAARALPLTSLTSAERPGQMLGGLPV 360  
DB 309 LGGGVAANLGRAASVGSLSVPOAAAQAVTPPAARALPLTSLTSAERPGQMLGGLPV 368  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
DB 369 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 399

RESULT 7  
US-09-287-849-22  
Sequence 22, Application US/09287849  
Patent No. US2002009459A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dilson, Davin C.  
APPLICANT: Alderson, Mark  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/09/287,849  
CURRENT FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 600  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
US-09-287-849-22

Query Match 100.0%; Score 1949; DB 9; Length 600;  
Best Local Similarity 100.0%; Pred. No. 1.2e-145;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVSADLFSASAFOSVVMGLTVGSMIG 60  
DB 9 MVDGALPPEINSAARMYAGPGSASLVAAQWMDVSADLFSASAFOSVVMGLTVGSMIG 68  
QY 61 SSAGLWVAASPVAWMSVTAGAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
DB 69 SSAGLWVAASPVAWMSVTAGAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMT 128



```

QY 121 LIATNLGONTPAIAVNEAEYGENMAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 180
|
|
|
Db 129 LIATNLGONTPAIAVNEAEYGENMAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 188
|
|
|
QY 181 LLEQAAAVEASDTAAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPHSRPSIN 240
|
|
|
Db 189 LLEQAAAVEASDTAAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPHSRPSIN 248
|
|
|
QY 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300
|
|
|
Db 249 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 308
|
|
|
QY 301 LGGGVAANLGRAASVGSLSVPOAAMAAANOAVTPAARALPLTSLTSAERPGOMLGGLPV 360
|
|
|
Db 309 LGGGVAANLGRAASVGSLSVPOAAMAAANOAVTPAARALPLTSLTSAERPGOMLGGLPV 368
|
|
|
QY 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAAG 391
|
|
|
Db 369 GOMGARAGGGLSGVLRVPPRPVMPHSPAAG 399
|
|
|

```

## RESULT 8

```

US-10-359-460-22
; Sequence 22, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-10-359-460-22

```

```

Query Match 100.0%; Score 1949; DB 12; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MWDGALPPEINSARNYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVMGLTVGSGWG 60
|
|
|
Db 9 MWDGALPPEINSARNYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVMGLTVGSGWG 68
|
|
|
QY 61 SSAGLWVAASPYVAMSVTAQAELTAAQVRYAAAYETAAGLTVPPVIAENRAELMI 120
|
|
|
Db 69 SSAGLWVAASPYVAMSVTAQAELTAAQVRYAAAYETAAGLTVPPVIAENRAELMI 128
|
|
|
QY 121 LIATNLGONTPAIAVNEAEYGENMAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 180
|
|
|

```

```

Db 129 LIATNLGONTPAIAVNEAEYGENMAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 188
|
|
|
QY 181 LLEQAAAVEASDTAAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPHSRPSIN 240
|
|
|
Db 189 LLEQAAAVEASDTAAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPHSRPSIN 248
|
|
|
QY 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300
|
|
|
Db 249 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 308
|
|
|
QY 301 LGGGVAANLGRAASVGSLSVPOAAMAAANOAVTPAARALPLTSLTSAERPGOMLGGLPV 360
|
|
|
Db 309 LGGGVAANLGRAASVGSLSVPOAAMAAANOAVTPAARALPLTSLTSAERPGOMLGGLPV 368
|
|
|
QY 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAAG 391
|
|
|
Db 369 GOMGARAGGGLSGVLRVPPRPVMPHSPAAG 399
|
|
|

```

## RESULT 9

```

US-10-098-732A-18
; Sequence 18, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderman, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72FmutSA
US-10-098-732A-18

```

```

Query Match 100.0%; Score 1949; DB 12; Length 729;
Best Local Similarity 100.0%; Pred. No. 1.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MWDGALPPEINSARNYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVMGLTVGSGWG 60
|
|
|
Db 142 MWDGALPPEINSARNYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVMGLTVGSGWG 201
|
|
|
QY 61 SSAGLWVAASPYVAMSVTAQAELTAAQVRYAAAYETAAGLTVPPVIAENRAELMI 120
|
|
|
Db 202 SSAGLWVAASPYVAMSVTAQAELTAAQVRYAAAYETAAGLTVPPVIAENRAELMI 261
|
|
|
QY 121 LIATNLGONTPAIAVNEAEYGENMAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 180
|
|
|
Db 262 LIATNLGONTPAIAVNEAEYGENMAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 321
|
|
|
QY 181 LLEQAAAVEASDTAAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPHSRPSIN 240
|
|
|
Db 322 LLEQAAAVEASDTAAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPHSRPSIN 381
|
|
|
QY 241 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300
|
|
|
Db 382 MWSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 441
|
|
|
QY 301 LGGGVAANLGRAASVGSLSVPOAAMAAANOAVTPAARALPLTSLTSAERPGOMLGGLPV 360
|
|
|

```

Db 442 LGGVAAANLGRAASVGLSVFQAMAAANQAVTPAARALPLTSLTSAERPGQMLGLPLV 501  
Qy 361 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAG 391  
Db 502 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAG 532

RESULT 10  
US-10-098-732A-65

Sequence 65, Application US/10098732A  
Publication No. US20030175294A1  
GENERAL INFORMATION:  
APPLICANT: Skeiky, Yasir  
APPLICANT: Brannon, Mark  
APPLICANT: Guderman, Jeffrey  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
FILE REFERENCE: 014058-012010US  
CURRENT APPLICATION NUMBER: US/10/098,732A  
PRIOR FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: US 60/275,837  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 65  
LENGTH: 930  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: MTB72F-MAPS  
OTHER INFORMATION: (T95F) fusion construct, TB MTB72F (Rat2-TbH9-Ra35)  
OTHER INFORMATION: linked to Leishmania thiol-specific antioxidant  
OTHER INFORMATION: (TSA or MAPS)  
US-10-098-732A-65

Query Match 100.0%; Score 1949; DB 12; Length 930;  
Best Local Similarity 100.0%; Pred. No. 2,2e-145;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDGALPPEINSAARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
Db 142 MVDGALPPEINSAARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 201  
Qy 61 SSAGLWAAASPVYVWMSVTAGAEELTAQVRVAAAAYETAYGTLVPPVIAENRAELMT 120  
Db 202 SSAGLWAAASPVYVWMSVTAGAEELTAQVRVAAAAYETAYGTLVPPVIAENRAELMT 261  
Qy 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 180  
Db 262 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 321  
Qy 181 LLEQAAVEBASDTAAANQIMNNVPOALQOAPTOGTTSSKLGGLMKTVPHRSPTSN 240  
Db 322 LLEQAAVEBASDTAAANQIMNNVPOALQOAPTOGTTSSKLGGLMKTVPHRSPTSN 381  
Qy 241 MVSMAANNHSMNTNSGVSMNTTSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSSG 300  
Db 382 MVSMAANNHSMNTNSGVSMNTTSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSSG 441  
Qy 301 LGGVAAANLGRAASVGLSVFQAMAAANQAVTPAARALPLTSLTSAERPGQMLGLPLV 360  
Db 442 LGGVAAANLGRAASVGLSVFQAMAAANQAVTPAARALPLTSLTSAERPGQMLGLPLV 501  
Qy 361 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAG 391  
Db 502 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAG 532

RESULT 11  
US-09-287-849-2  
Sequence 2, Application US/09287849  
Patent No. US2002009459A1

GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Alderson, Mark  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/09/287,849  
CURRENT FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 729  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
US-09-287-849-2

Query Match 99.7%; Score 1944; DB 9; Length 729;  
Best Local Similarity 99.7%; Pred. No. 3,9e-145;  
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVDGALPPEINSAARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
Db 142 MVDGALPPEINSAARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 201  
Qy 61 SSAGLWAAASPVYVWMSVTAGAEELTAQVRVAAAAYETAYGTLVPPVIAENRAELMT 120  
Db 202 SSAGLWAAASPVYVWMSVTAGAEELTAQVRVAAAAYETAYGTLVPPVIAENRAELMT 261  
Qy 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 180  
Db 262 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 321  
Qy 181 LLEQAAVEBASDTAAANQIMNNVPOALQOAPTOGTTSSKLGGLMKTVPHRSPTSN 240  
Db 322 LLEQAAVEBASDTAAANQIMNNVPOALQOAPTOGTTSSKLGGLMKTVPHRSPTSN 381  
Qy 241 MVSMAANNHSMNTNSGVSMNTTSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSSG 300  
Db 382 MVSMAANNHSMNTNSGVSMNTTSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSSG 441  
Qy 301 LGGVAAANLGRAASVGLSVFQAMAAANQAVTPAARALPLTSLTSAERPGQMLGLPLV 360  
Db 442 LGGVAAANLGRAASVGLSVFQAMAAANQAVTPAARALPLTSLTSAERPGQMLGLPLV 501  
Qy 361 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAG 391  
Db 502 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAG 532

RESULT 12  
US-10-359-460-2  
Sequence 2, Application US/10359460  
Publication No. US20030147911A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.

```

; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-00902005
; CURRENT APPLICATION NUMBER: US/10/359,460
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; US-10-359,460-2

```

Query Match Best Local Similarity 99.7%; Score 1944; DB 12; Length 729; Pred. No. 3,9e-145;

Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSAIDFSAASAFQSVVMGLTVGWSWG 60
DB 142 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSAIDFSAASAFQSVVMGLTVGWSWG 201
QY 61 SSAGLVAAASPYVAMSVTAQAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLVAAASPYVAMSVTAQAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261
QY 121 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPPEAPEMTSAGG 180
DB 262 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPPEAPEMTSAGG 321
QY 181 LLEQAAAVEASDTAAANOQMNVPALQOLAQPTGTTSSKLGGLMKTVPHSRPSIN 240
DB 322 LLEQAAAVEASDTAAANOQMNVPALQOLAQPTGTTSSKLGGLMKTVPHSRPSIN 381
QY 241 MYSMANNHSMNTNSGVSMNTTLLSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSSG 300
DB 382 MYSMANNHSMNTNSGVSMNTTLLSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSSG 441
QY 301 LGGGVAAANIGRAASVGSLSVPOAMAAANOAVTPARALPLTSLTSAERGPGOMLGGLPV 360
DB 442 LGGGVAAANIGRAASVGSLSVPOAMAAANOAVTPARALPLTSLTSAERGPGOMLGGLPV 501
QY 361 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAAG 391
DB 502 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAAG 532

```

```

; RESULT 13
; US-10-098-732A-16
; Sequence 16, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderman, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a

```

```

; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-01201005
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; OTHER INFORMATION: protein MTB72F (Ral2-TBh9-Ra35 or MTB32-MTB39
; OTHER INFORMATION: fusion)
; US-10-098-732A-16

```

Query Match Best Local Similarity 99.7%; Score 1944; DB 12; Length 729; Pred. No. 3,9e-145;

Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSAIDFSAASAFQSVVMGLTVGWSWG 60
DB 142 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSAIDFSAASAFQSVVMGLTVGWSWG 201
QY 61 SSAGLVAAASPYVAMSVTAQAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLVAAASPYVAMSVTAQAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261
QY 121 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPPEAPEMTSAGG 180
DB 262 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPPEAPEMTSAGG 321
QY 181 LLEQAAAVEASDTAAANOQMNVPALQOLAQPTGTTSSKLGGLMKTVPHSRPSIN 240
DB 322 LLEQAAAVEASDTAAANOQMNVPALQOLAQPTGTTSSKLGGLMKTVPHSRPSIN 381
QY 241 MYSMANNHSMNTNSGVSMNTTLLSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSSG 300
DB 382 MYSMANNHSMNTNSGVSMNTTLLSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSSG 441
QY 301 LGGGVAAANIGRAASVGSLSVPOAMAAANOAVTPARALPLTSLTSAERGPGOMLGGLPV 360
DB 442 LGGGVAAANIGRAASVGSLSVPOAMAAANOAVTPARALPLTSLTSAERGPGOMLGGLPV 501
QY 361 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAAG 391
DB 502 GQMGARAGGGLSGVLRVPPRPVYVMPHSPAAAG 532

```

```

; RESULT 14
; US-10-084-843-111
; Sequence 11, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, David C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Iwardik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington

```

```

; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 111:
US-10-084-843-111

```

```

Query Match      84.8%; Score 1652.5; DB 12; Length 396;
Best Local Similarity 84.9%; Pred. No. 1.9e-122;
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPEINSAARMYAGSGSASLVAAQOMDSVASDLFSAASAFOSVVMGLTVGSWIG 60
DB 1 VVDFGALPEINSAARMYAGSGSASLVAAQOMDSVASDLFSAASAFOSVVMGLTVGSWIG 60
QY 61 SSAGLMTVAASPYVAMSVTAAGAEITPAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120
DB 61 SSAGLMTVAASPYVAMSVTAAGAEITPAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120
QY 121 LIATNLGQONTPAIAVNEAEYGEWMAODAAAMFGYAAATATATATLTPFEEDAPMTSAG 180
DB 121 LIATNLGQONTPAIAVNEAEYGEWMAODAAAMFGYAAATATATATLTPFEEDAPMTSAG 180
QY 181 LLEQAAAVEASDPAANQOLMNNVPOALQOLAPPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 181 LLEQAAAVEASDPAANQOLMNNVPOALQOLAPPTGTTSSKLGIMKTVPSPHRSPISN 240
QY 241 MVSANNNHSMSTNGSVMTNTLSMTLKGFPAPAAAQAQVTAONGVRAMSS---LGSSL 296
DB 241 IVSMLNHHVSMSTNGSVMASTLHSMKGFAP-AAAQAVETAAQNGVQAMSSLGSL 299
QY 297 GSSGLGGVAAANTLGRAASVGSLSVPOAWAANAQAVTPPARALPLTSLTSAERPGQMTG 356
DB 300 GSSGLGAGVAAANTLGRAASVGSLSVPOAWAANAQAVTPPARALPLTSLTSAERPGHMLG 359
QY 357 GLPVGMGARAG--GGLSGVLRVPPRPVYMPHSPAAG 391
DB 360 GLPLGQLTNSGGGFGVSNALRMPPRAYVMPRVPAAG 396

```

```

RESULT 15
US-10-193-002-106
; Sequence 106, Application US/10193002
; Publication No. US20030135026a1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.M.
; Dillon, Davin C.
; Campos-Neco, Antonia
; Houghton, Raymond

```

```

; Vedick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-10-193-002-106

```

```

Query Match      84.8%; Score 1652.5; DB 12; Length 396;
Best Local Similarity 84.9%; Pred. No. 1.9e-122;
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPEINSAARMYAGSGSASLVAAQOMDSVASDLFSAASAFOSVVMGLTVGSWIG 60
DB 1 VVDFGALPEINSAARMYAGSGSASLVAAQOMDSVASDLFSAASAFOSVVMGLTVGSWIG 60
QY 61 SSAGLMTVAASPYVAMSVTAAGAEITPAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120
DB 61 SSAGLMTVAASPYVAMSVTAAGAEITPAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120
QY 121 LIATNLGQONTPAIAVNEAEYGEWMAODAAAMFGYAAATATATATLTPFEEDAPMTSAG 180
DB 121 LIATNLGQONTPAIAVNEAEYGEWMAODAAAMFGYAAATATATATLTPFEEDAPMTSAG 180
QY 181 LLEQAAAVEASDPAANQOLMNNVPOALQOLAPPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 181 LLEQAAAVEASDPAANQOLMNNVPOALQOLAPPTGTTSSKLGIMKTVPSPHRSPISN 240
QY 241 MVSANNNHSMSTNGSVMTNTLSMTLKGFPAPAAAQAQVTAONGVRAMSS---LGSSL 296
DB 241 IVSMLNHHVSMSTNGSVMASTLHSMKGFAP-AAAQAVETAAQNGVQAMSSLGSL 299
QY 297 GSSGLGGVAAANTLGRAASVGSLSVPOAWAANAQAVTPPARALPLTSLTSAERPGQMTG 356
DB 300 GSSGLGAGVAAANTLGRAASVGSLSVPOAWAANAQAVTPPARALPLTSLTSAERPGHMLG 359
QY 357 GLPVGMGARAG--GGLSGVLRVPPRPVYMPHSPAAG 391
DB 360 GLPLGQLTNSGGGFGVSNALRMPPRAYVMPRVPAAG 396

```

Tue Nov 25 10:11:31 2003

us-09-724-685-107.rapb

Page 9

Search completed: November 21, 2003, 17:11:03  
Job time : 34 secs

---

**THIS PAGE BLANK (USPTO)**



```
QY 301 LGGVAAANLGRAASVGLSVPOAAANQAVTPARALPLTSLTSAERGPQMLGGLPV 360
|
|
|
Db 301 LGGGVAAANLGRAASVGLSVPOAAANQAVTPARALPLTSLTSAERGPQMLGGLPV 360
QY 361 GQMGARAGGGLSGVLRRVPPRPVYVPHSPAA 391
|
|
|
Db 361 GQMGARAGGGLSGVLRRVPPRPVYVPHSPAA 391

RESULT 2
US-09-886-349A-20
; Sequence 20, Application US/09886349A
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 20
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: bi-fusion
; OTHER INFORMATION: protein TBH9-Ra35 (designated MTB59F)
; US-09-886-349A-20

Query Match 100.0%; Score 1949; DB 5; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.3e-108;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINARMYAGGSAISLVAAQMDVSASDLFSASAFQSVVWGLTVGSMIG 60
|
|
|
Db 9 MVDGALPPEINARMYAGGSAISLVAAQMDVSASDLFSASAFQSVVWGLTVGSMIG 68
QY 61 SSAGLWVAASPYVAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
|
|
|
Db 69 SSAGLWVAASPYVAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 128
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGVAAATATATATLLPFEAPDMTSAGG 180
|
|
|
Db 129 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGVAAATATATATLLPFEAPDMTSAGG 188
QY 181 LLEQAAVEBASDTAAANQLMNNVPAALQOLAQPTGTTSSKLGIMKTVPSPHRSPI 240
|
|
|
Db 189 LLEQAAVEBASDTAAANQLMNNVPAALQOLAQPTGTTSSKLGIMKTVPSPHRSPI 248
QY 241 MVSMAHHMSTMTSGVSMNTTSSMLKGFAPAPAAAQAVQTPAONGVAMSSLSGSSG 300
|
|
|
Db 249 MVSMAHHMSTMTSGVSMNTTSSMLKGFAPAPAAAQAVQTPAONGVAMSSLSGSSG 308
QY 301 LGGGVAAANLGRAASVGLSVPOAAANQAVTPARALPLTSLTSAERGPQMLGGLPV 360
|
|
|
Db 309 LGGGVAAANLGRAASVGLSVPOAAANQAVTPARALPLTSLTSAERGPQMLGGLPV 368
QY 361 GQMGARAGGGLSGVLRRVPPRPVYVPHSPAA 391
|
|
|
Db 369 GQMGARAGGGLSGVLRRVPPRPVYVPHSPAA 399

RESULT 3
US-10-369-983-2
; Sequence 2, Application US/10369983
```

```
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guadian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mutated
; OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32mutSA)
; US-10-369-983-2

Query Match 100.0%; Score 1949; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINARMYAGGSAISLVAAQMDVSASDLFSASAFQSVVWGLTVGSMIG 60
|
|
|
Db 333 MVDGALPPEINARMYAGGSAISLVAAQMDVSASDLFSASAFQSVVWGLTVGSMIG 392
QY 61 SSAGLWVAASPYVAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
|
|
|
Db 393 SSAGLWVAASPYVAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 452
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGVAAATATATATLLPFEAPDMTSAGG 180
|
|
|
Db 453 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGVAAATATATATLLPFEAPDMTSAGG 512
QY 181 LLEQAAVEBASDTAAANQLMNNVPAALQOLAQPTGTTSSKLGIMKTVPSPHRSPI 240
|
|
|
Db 513 LLEQAAVEBASDTAAANQLMNNVPAALQOLAQPTGTTSSKLGIMKTVPSPHRSPI 572
QY 241 MVSMAHHMSTMTSGVSMNTTSSMLKGFAPAPAAAQAVQTPAONGVAMSSLSGSSG 300
|
|
|
Db 573 MVSMAHHMSTMTSGVSMNTTSSMLKGFAPAPAAAQAVQTPAONGVAMSSLSGSSG 632
QY 301 LGGGVAAANLGRAASVGLSVPOAAANQAVTPARALPLTSLTSAERGPQMLGGLPV 360
|
|
|
Db 633 LGGGVAAANLGRAASVGLSVPOAAANQAVTPARALPLTSLTSAERGPQMLGGLPV 692
QY 361 GQMGARAGGGLSGVLRRVPPRPVYVPHSPAA 391
|
|
|
Db 693 GQMGARAGGGLSGVLRRVPPRPVYVPHSPAA 723

RESULT 4
US-09-886-349A-18
; Sequence 18, Application US/09886349A
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.1
```



```

; SEQ ID NO 18
; LENGTH: 729
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72FmutSA
US-09-886-349A-18

Query Match      100.0%; Score 1949; DB 5; Length 729;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60
DB 142 MVDGALPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 201

QY 61 SSAGLMVAAAAPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLMVAAAAPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261

QY 121 LIATNLGONTPALIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180
DB 262 LIATNLGONTPALIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 321

QY 181 LLEQAAVEASDPTAAANQNMNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 322 LLEQAAVEASDPTAAANQNMNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 381

QY 241 MWSMANNHSMNTSGVSMNTLTSSMLKGFAPAAAQAQVOTAQNGVRAMSLGSSIGSSG 300
DB 382 MWSMANNHSMNTSGVSMNTLTSSMLKGFAPAAAQAQVOTAQNGVRAMSLGSSIGSSG 441

QY 301 LGGVVAANLGRAASVGSLSVPQAMAAANQVTPAARALPLTSLTSAERPGQMLGCLPV 360
DB 442 LGGVVAANLGRAASVGSLSVPQAMAAANQVTPAARALPLTSLTSAERPGQMLGCLPV 501

QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 391
DB 502 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 532

RESULT 5
US-10-369-983-21
; Sequence 21, Application US/10369983
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 729
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72F
US-10-369-983-21

Query Match      100.0%; Score 1949; DB 6; Length 729;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60
DB 142 MVDGALPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 201
```

```

QY 61 SSAGLMVAAAAPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLMVAAAAPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261

QY 121 LIATNLGONTPALIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180
DB 262 LIATNLGONTPALIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 321

QY 181 LLEQAAVEASDPTAAANQNMNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 322 LLEQAAVEASDPTAAANQNMNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 381

QY 241 MWSMANNHSMNTSGVSMNTLTSSMLKGFAPAAAQAQVOTAQNGVRAMSLGSSIGSSG 300
DB 382 MWSMANNHSMNTSGVSMNTLTSSMLKGFAPAAAQAQVOTAQNGVRAMSLGSSIGSSG 441

QY 301 LGGVVAANLGRAASVGSLSVPQAMAAANQVTPAARALPLTSLTSAERPGQMLGCLPV 360
DB 442 LGGVVAANLGRAASVGSLSVPQAMAAANQVTPAARALPLTSLTSAERPGQMLGCLPV 501

QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 391
DB 502 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 532

RESULT 6
US-10-369-983-22
; Sequence 22, Application US/10369983
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 22
; LENGTH: 729
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated
US-10-369-983-22

Query Match      100.0%; Score 1949; DB 6; Length 729;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60
DB 142 MVDGALPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 201

QY 61 SSAGLMVAAAAPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLMVAAAAPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261

QY 121 LIATNLGONTPALIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180
DB 262 LIATNLGONTPALIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 321

QY 181 LLEQAAVEASDPTAAANQNMNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 322 LLEQAAVEASDPTAAANQNMNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 381

QY 241 MWSMANNHSMNTSGVSMNTLTSSMLKGFAPAAAQAQVOTAQNGVRAMSLGSSIGSSG 300
```

Db 382 MVSAMNNHSMNTSGVSMNTTLLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 441  
 QY 301 LGGVAAANTGRASVGSLSVPOAAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
 Db 442 LGGVAAANTGRASVGSLSVPOAAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 501  
 QY 361 GOMGARAGGGLSGVLRAVPPRPYVMPHSPAAG 391  
 Db 502 GOMGARAGGGLSGVLRAVPPRPYVMPHSPAAG 532

## RESULT 7

US-10-369-983-15  
 ; Sequence 15, Application US/10369983  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Guderman, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-009081US  
 ; CURRENT APPLICATION NUMBER: US/10/369,983  
 ; CURRENT FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; PRIOR FILING DATE: 2002-02-15  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 15  
 ; LENGTH: 813  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; US-10-369-983-15

Query Match 100.0%; Score 1949; DB 6; Length 813;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-108;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGSMIG 60  
 Db 142 MVDGALPPEINSAARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGSMIG 201  
 QY 61 SSAGLWVAASPVAAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
 Db 202 SSAGLWVAASPVAAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 261  
 QY 121 LIATNLLGONTPTAIAVNEAEYGEWMAQDAAAMFGYAAATATATATLLPFEAEPMTSAGG 180  
 Db 262 LIATNLLGONTPTAIAVNEAEYGEWMAQDAAAMFGYAAATATATATLLPFEAEPMTSAGG 321  
 QY 181 LLEQAAAVEASPTAAANQLMNNVPOALQOLAOPTGTPSSKLGIMKTIVSPHRSPI 240  
 Db 322 LLEQAAAVEASPTAAANQLMNNVPOALQOLAOPTGTPSSKLGIMKTIVSPHRSPI 381  
 QY 241 MVSAMNNHSMNTSGVSMNTTLLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 300  
 Db 382 MVSAMNNHSMNTSGVSMNTTLLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 441  
 QY 301 LGGVAAANTGRASVGSLSVPOAAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
 Db 442 LGGVAAANTGRASVGSLSVPOAAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 501  
 QY 361 GOMGARAGGGLSGVLRAVPPRPYVMPHSPAAG 391  
 Db 502 GOMGARAGGGLSGVLRAVPPRPYVMPHSPAAG 532

## RESULT 8

US-10-369-983-14  
 ; Sequence 14, Application US/10369983  
 ; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Guderman, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-009081US  
 ; CURRENT APPLICATION NUMBER: US/10/369,983  
 ; CURRENT FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; PRIOR FILING DATE: 2002-02-15  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 825  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; US-10-369-983-14

Query Match 100.0%; Score 1949; DB 6; Length 825;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-108;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSAARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGSMIG 60  
 Db 142 MVDGALPPEINSAARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVMGLTVGSMIG 201  
 QY 61 SSAGLWVAASPVAAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
 Db 202 SSAGLWVAASPVAAMSVTAAGAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 261  
 QY 121 LIATNLLGONTPTAIAVNEAEYGEWMAQDAAAMFGYAAATATATATLLPFEAEPMTSAGG 180  
 Db 262 LIATNLLGONTPTAIAVNEAEYGEWMAQDAAAMFGYAAATATATATLLPFEAEPMTSAGG 321  
 QY 181 LLEQAAAVEASPTAAANQLMNNVPOALQOLAOPTGTPSSKLGIMKTIVSPHRSPI 240  
 Db 322 LLEQAAAVEASPTAAANQLMNNVPOALQOLAOPTGTPSSKLGIMKTIVSPHRSPI 381  
 QY 241 MVSAMNNHSMNTSGVSMNTTLLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 300  
 Db 382 MVSAMNNHSMNTSGVSMNTTLLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 441  
 QY 301 LGGVAAANTGRASVGSLSVPOAAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
 Db 442 LGGVAAANTGRASVGSLSVPOAAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 501  
 QY 361 GOMGARAGGGLSGVLRAVPPRPYVMPHSPAAG 391  
 Db 502 GOMGARAGGGLSGVLRAVPPRPYVMPHSPAAG 532

## RESULT 9

US-10-369-983-13  
 ; Sequence 13, Application US/10369983  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Guderman, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-009081US  
 ; CURRENT APPLICATION NUMBER: US/10/369,983  
 ; CURRENT FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; PRIOR FILING DATE: 2002-02-15  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 13  
 ; LENGTH: 875  
 ; TYPE: PRT

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: MTB89F (MTB72F-Erd14)
US-10-369-983-13

Query Match
  Best Local Similarity 100.0%; Score 1949; DB 6; Length 875;
  Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARYAGSGASLVAAQWMDVSASDLFSAASAFQSVVWGLTVGSWIG 60
DB 142 MVDGALPEINSAARYAGSGASLVAAQWMDVSASDLFSAASAFQSVVWGLTVGSWIG 201
QY 61 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 261
QY 121 LIATNLLGONTPTAIAVNEAEYGEWMAODAAAFGYAATAATATATLLPFEAEPMTSAGG 180
DB 262 LIATNLLGONTPTAIAVNEAEYGEWMAODAAAFGYAATAATATATLLPFEAEPMTSAGG 321
QY 181 LLEQAAAVEASPTAANQLMNNVPOALQOLAOPTGTTSSKLGLMKTVPSPHRSPISN 240
DB 322 LLEQAAAVEASPTAANQLMNNVPOALQOLAOPTGTTSSKLGLMKTVPSPHRSPISN 381
QY 241 MWSMANNHSMNTSGVSMNTTLLSMKGFAPAAAQAVQTAQNGVRAMSLGSSSG 300
DB 382 MWSMANNHSMNTSGVSMNTTLLSMKGFAPAAAQAVQTAQNGVRAMSLGSSSG 441
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANOAVTPPAALPLTSLTSAERPGQMLGGLPV 360
DB 442 LGGGVAANLGRAASVGSLSVPOMAAANOAVTPPAALPLTSLTSAERPGQMLGGLPV 501
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391
DB 502 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 532

RESULT 10
US-10-369-983-12
; Sequence 12, Application US/10369983
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderman, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369, 983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-12

Query Match
  Best Local Similarity 100.0%; Score 1949; DB 6; Length 930;
  Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARYAGSGASLVAAQWMDVSASDLFSAASAFQSVVWGLTVGSWIG 60
DB 142 MVDGALPEINSAARYAGSGASLVAAQWMDVSASDLFSAASAFQSVVWGLTVGSWIG 201
QY 61 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 202 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 120
```

```

DB 202 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 261
QY 121 LIATNLLGONTPTAIAVNEAEYGEWMAODAAAFGYAATAATATATLLPFEAEPMTSAGG 180
DB 262 LIATNLLGONTPTAIAVNEAEYGEWMAODAAAFGYAATAATATATATLLPFEAEPMTSAGG 321
QY 181 LLEQAAAVEASPTAANQLMNNVPOALQOLAOPTGTTSSKLGLMKTVPSPHRSPISN 240
DB 322 LLEQAAAVEASPTAANQLMNNVPOALQOLAOPTGTTSSKLGLMKTVPSPHRSPISN 381
QY 241 MWSMANNHSMNTSGVSMNTTLLSMKGFAPAAAQAVQTAQNGVRAMSLGSSSG 300
DB 382 MWSMANNHSMNTSGVSMNTTLLSMKGFAPAAAQAVQTAQNGVRAMSLGSSSG 441
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANOAVTPPAALPLTSLTSAERPGQMLGGLPV 360
DB 442 LGGGVAANLGRAASVGSLSVPOMAAANOAVTPPAALPLTSLTSAERPGQMLGGLPV 501
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391
DB 502 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 532

RESULT 11
US-10-369-983-4
; Sequence 4, Application US/10369983
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderman, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369, 983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MTB-102F fusion
; OTHER INFORMATION: protein
US-10-369-983-4

Query Match
  Best Local Similarity 100.0%; Score 1949; DB 6; Length 1010;
  Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARYAGSGASLVAAQWMDVSASDLFSAASAFQSVVWGLTVGSWIG 60
DB 333 MVDGALPEINSAARYAGSGASLVAAQWMDVSASDLFSAASAFQSVVWGLTVGSWIG 332
QY 61 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 393 SSAGLWVAASPYVAMSVTAAGAEELTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 452
QY 121 LIATNLLGONTPTAIAVNEAEYGEWMAODAAAFGYAATAATATATLLPFEAEPMTSAGG 180
DB 453 LIATNLLGONTPTAIAVNEAEYGEWMAODAAAFGYAATAATATATATLLPFEAEPMTSAGG 512
QY 181 LLEQAAAVEASPTAANQLMNNVPOALQOLAOPTGTTSSKLGLMKTVPSPHRSPISN 240
DB 513 LLEQAAAVEASPTAANQLMNNVPOALQOLAOPTGTTSSKLGLMKTVPSPHRSPISN 572
QY 241 MWSMANNHSMNTSGVSMNTTLLSMKGFAPAAAQAVQTAQNGVRAMSLGSSSG 300
DB 573 MWSMANNHSMNTSGVSMNTTLLSMKGFAPAAAQAVQTAQNGVRAMSLGSSSG 632
```

QY 301 LGGVAAANLGRAASVGSLSVPOAAMAAQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
DB 633 LGGVAAANLGRAASVGSLSVPOAAMAAQAVTPAARALPLTSLTSAERGGQMLGGLPV 692  
QY 361 GOMGARAGGGLSGVLRVPPRPYVPHSPAG 391  
DB 693 GOMGARAGGGLSGVLRVPPRPYVPHSPAG 723

## RESULT 12

US-10-369-983-18  
; Sequence 18, Application US/10369983  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderman, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; CURRENT FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 1016  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
; OTHER INFORMATION: MTB103F (MTB72F-85b)  
US-10-369-983-18

Query Match 100.0%; Score 1949; DB 6; Length 1016;

Best Local Similarity 100.0%; Pred. No. 2.4e-108; Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSGARMYAGPGSASLVAAAQWMDVASDLFSAASAFQSVWGLTVGSWIG 60  
DB 142 MVDGALPPEINSGARMYAGPGSASLVAAAQWMDVASDLFSAASAFQSVWGLTVGSWIG 201  
QY 61 SSAGLWVAASPYVAMWSVTAQAEELTAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 202 SSAGLWVAASPYVAMWSVTAQAEELTAQVRVAAAAYETAYGLTVPPVIAENRAELMI 261  
QY 121 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGAAATATATATLLPFEEAPEMTSAGG 180  
DB 262 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGAAATATATATLLPFEEAPEMTSAGG 321  
QY 181 LLEQAAAVEASDTAAANQMLNNVPQALQOLAOPTGTTSSKLGLMKTVPSPRSPISN 240  
DB 322 LLEQAAAVEASDTAAANQMLNNVPQALQOLAOPTGTTSSKLGLMKTVPSPRSPISN 381  
QY 241 MVSNNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVOTAONGVRAMSSLGSSSG 300  
DB 382 MVSNNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVOTAONGVRAMSSLGSSSG 441  
QY 301 LGGVAAANLGRAASVGSLSVPOAAMAAQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
DB 442 LGGVAAANLGRAASVGSLSVPOAAMAAQAVTPAARALPLTSLTSAERGGQMLGGLPV 501  
QY 361 GOMGARAGGGLSGVLRVPPRPYVPHSPAG 391  
DB 502 GOMGARAGGGLSGVLRVPPRPYVPHSPAG 532

RESULT 13  
US-10-369-983-17  
; Sequence 17, Application US/10369983  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderman, Jeff

; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; CURRENT FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 1022  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
; OTHER INFORMATION: MTB102cm2F (MTB72F-hTCC#1)  
US-10-369-983-17

Query Match 100.0%; Score 1949; DB 6; Length 1022;

Best Local Similarity 100.0%; Pred. No. 2.4e-108; Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSGARMYAGPGSASLVAAAQWMDVASDLFSAASAFQSVWGLTVGSWIG 60  
DB 142 MVDGALPPEINSGARMYAGPGSASLVAAAQWMDVASDLFSAASAFQSVWGLTVGSWIG 201  
QY 61 SSAGLWVAASPYVAMWSVTAQAEELTAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 202 SSAGLWVAASPYVAMWSVTAQAEELTAQVRVAAAAYETAYGLTVPPVIAENRAELMI 261  
QY 121 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGAAATATATATLLPFEEAPEMTSAGG 180  
DB 262 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGAAATATATATLLPFEEAPEMTSAGG 321  
QY 181 LLEQAAAVEASDTAAANQMLNNVPQALQOLAOPTGTTSSKLGLMKTVPSPRSPISN 240  
DB 322 LLEQAAAVEASDTAAANQMLNNVPQALQOLAOPTGTTSSKLGLMKTVPSPRSPISN 381  
QY 241 MVSNNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVOTAONGVRAMSSLGSSSG 300  
DB 382 MVSNNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVOTAONGVRAMSSLGSSSG 441  
QY 301 LGGVAAANLGRAASVGSLSVPOAAMAAQAVTPAARALPLTSLTSAERGGQMLGGLPV 360  
DB 442 LGGVAAANLGRAASVGSLSVPOAAMAAQAVTPAARALPLTSLTSAERGGQMLGGLPV 501  
QY 361 GOMGARAGGGLSGVLRVPPRPYVPHSPAG 391  
DB 502 GOMGARAGGGLSGVLRVPPRPYVPHSPAG 532

## RESULT 14

US-10-369-983-16  
; Sequence 16, Application US/10369983  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderman, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; CURRENT FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 1154  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 OTHER INFORMATION: MTB172F (MTB72F-MTCH2)  
 US-10-369-983-16

Query Match 100.0%; Score 1949; DB 6; Length 1154;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-108;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEPINSARVAYAGSGASLVAAAQMDVSADLFSASAFQSVYVWGLTVGSMWG 60  
 DB 142 MVDGALPEPINSARVAYAGSGASLVAAAQMDVSADLFSASAFQSVYVWGLTVGSMWG 201  
 QY 61 SSAGLWVAASPVMMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
 DB 202 SSAGLWVAASPVMMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261  
 QY 121 LIATNLGONTPAIAVNEAEYGEEMAQDAAMFGVAAATATATATLLPFEAEPEMTSAGG 180  
 DB 262 LIATNLGONTPAIAVNEAEYGEEMAQDAAMFGVAAATATATATLLPFEAEPEMTSAGG 321  
 QY 181 LLEQAAVEASDPTAAANQMLNNVPOALQOAPTOGTTPSSKLGIMKTVSFHRSPISN 240  
 DB 322 LLEQAAVEASDPTAAANQMLNNVPOALQOAPTOGTTPSSKLGIMKTVSFHRSPISN 381  
 QY 241 MYSMANNHSMNTNSGVSMNTTSSMLKGFAPAAAAQAVOTAAQNGVRAMSSLGSSSG 300  
 DB 382 MYSMANNHSMNTNSGVSMNTTSSMLKGFAPAAAAQAVOTAAQNGVRAMSSLGSSSG 441  
 QY 301 LGGGVAANIGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
 DB 442 LGGGVAANIGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 501  
 QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
 DB 502 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 532

## RESULT 15

US-09-886-349A-16  
 ; Sequence 16, Application US/09886349A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Corix Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-00907005  
 ; CURRENT APPLICATION NUMBER: US/09/886,349A  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: US 09/597,796  
 ; PRIOR FILING DATE: 2000-06-20  
 ; PRIOR APPLICATION NUMBER: US 60/265,737  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 16:  
 ; LENGTH: 729  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
 ; OTHER INFORMATION: protein MTB72F (Ra12-TbH9-Ra35 or MTB32-MTB39  
 ; OTHER INFORMATION: fusion)  
 US-09-886-349A-16

Query Match 99.7%; Score 1944; DB 5; Length 729;  
 Best Local Similarity 99.7%; Pred. No. 3.3e-108;  
 Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVDGALPEPINSARVAYAGSGASLVAAAQMDVSADLFSASAFQSVYVWGLTVGSMWG 60  
 DB 142 MVDGALPEPINSARVAYAGSGASLVAAAQMDVSADLFSASAFQSVYVWGLTVGSMWG 201

QY 61 SSAGLWVAASPVMMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
 DB 202 SSAGLWVAASPVMMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261  
 QY 121 LIATNLGONTPAIAVNEAEYGEEMAQDAAMFGVAAATATATATLLPFEAEPEMTSAGG 180  
 DB 262 LIATNLGONTPAIAVNEAEYGEEMAQDAAMFGVAAATATATATLLPFEAEPEMTSAGG 321  
 QY 181 LLEQAAVEASDPTAAANQMLNNVPOALQOAPTOGTTPSSKLGIMKTVSFHRSPISN 240  
 DB 322 LLEQAAVEASDPTAAANQMLNNVPOALQOAPTOGTTPSSKLGIMKTVSFHRSPISN 381  
 QY 241 MYSMANNHSMNTNSGVSMNTTSSMLKGFAPAAAAQAVOTAAQNGVRAMSSLGSSSG 300  
 DB 382 MYSMANNHSMNTNSGVSMNTTSSMLKGFAPAAAAQAVOTAAQNGVRAMSSLGSSSG 441  
 QY 301 LGGGVAANIGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
 DB 442 LGGGVAANIGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 501  
 QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
 DB 502 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 532

Search completed: November 21, 2003, 17:10:17  
 Job time : 21 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 16:55:09 / Search time 45 Seconds  
(without alignments)  
1379.158 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949  
Sequence: 1 MVDGALPPEINSARMYAGP.....SGVLKVPYPPVMPHSPAAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: A.Geneseq.19Jun03.\*  
2: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1949	100.0	391	18	AAW32449
2	1949	100.0	391	18	AAW32381
3	1949	100.0	391	19	AAW81702
4	1949	100.0	391	19	AAW64335
5	1949	100.0	391	20	AAI39132
6	1949	100.0	391	20	AAI38989
7	1949	100.0	391	20	AAV04778
8	1949	100.0	391	22	AAU01888
9	1949	100.0	391	23	AAE29707

10	1949	100.0	391	23	AAE17571	Mycobacterium spec
11	1949	100.0	596	23	AAV32070	Mycobacterium tube
12	1949	100.0	596	23	AAE29710	Mycobacterium sp.
13	1949	100.0	596	23	AAE17574	Mycobacterium spec
14	1949	100.0	599	23	AAV74599	Antigenic fusion p
15	1949	100.0	600	20	AAV32068	Mycobacterium tube
16	1949	100.0	600	23	AAU74597	Antigenic fusion p
17	1949	100.0	729	22	AAO22142	Ra12-H9-32A fusion
18	1949	100.0	729	23	AAE29709	Mycobacterium sp.
19	1949	100.0	729	23	AAE17573	Mycobacterium spec
20	1949	100.0	744	22	AAU01902	M. tuberculosis an
21	1949	100.0	815	22	AAU01904	M. tuberculosis an
22	1949	100.0	930	23	AAE29731	Mycobacterium sp.
23	1949	99.8	788	22	AAU01903	M. tuberculosis an
24	1944	99.7	394	20	AAV04779	Mycobacterium spec
25	1944	99.7	729	23	AAE29708	Mycobacterium spec
26	1944	99.7	729	23	AAE17572	Mycobacterium spec
27	1931	99.1	729	20	AAV32059	Mycobacterium tube
28	1902.5	97.6	726	23	AAU74588	Antigenic fusion p
29	1652.5	84.8	396	19	AAW81704	M. tuberculosis im
30	1652.5	84.8	396	19	AAW64337	Mycobacterium tube
31	1652.5	84.8	396	20	AAV39134	M. tuberculosis an
32	1652.5	84.8	396	20	AAV38991	M. tuberculosis an
33	1486.5	76.3	359	19	AAW81703	M. tuberculosis re
34	1486.5	76.3	359	19	AAW64336	Mycobacterium tube
35	1486.5	76.3	359	20	AAV39133	M. tuberculosis re
36	1486.5	76.3	359	20	AAV38990	M. tuberculosis re
37	1474.5	75.7	341	18	AAW32451	Mycobacterium tube
38	1474.5	75.7	341	18	AAW32383	Mycobacterium tube
39	1187.5	60.9	358	18	AAU74591	Antigenic fusion p
40	1187	60.9	263	18	AAW32447	Mycobacterium tube
41	1187	60.9	263	18	AAW32479	Mycobacterium tube
42	1187	60.9	263	19	AAW81680	M. tuberculosis im
43	1187	60.9	263	19	AAW64317	Mycobacterium tube
44	1187	60.9	263	20	AAV32062	Mycobacterium tube
45	1187	60.9	263	20	AAV39121	M. tuberculosis an

#### ALIGNMENTS

RESULT 1  
AAW32449 standard; Protein; 391 AA.  
ID AAW32449;  
AC AAW32449;  
XX  
XX 09-JAN-1998 (first entry)  
DT  
XX  
DE Mycobacterium tuberculosis antigen Tbh-9FL.  
XX  
XX Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
KW skin testing; M.tuberculosis.  
XX  
OS Mycobacterium tuberculosis.  
XX  
XX  
PN WO9709428-A2.  
XX  
XX 13-MAR-1997.  
PD  
XX  
XX 30-AUG-1996; 96WO-US14674.  
PF  
XX  
PR 12-JUL-1996; 96US-0680574.  
PR 01-SEP-1995; 95US-0523436.  
PR 22-SEP-1995; 95US-053634.  
PR 22-MAR-1996; 96US-0620874.  
PR 05-JUN-1996; 96US-0659683.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Campos-neto A, Dillon DC, Houghton R, Reed SG, Skelky YAW;  
PI Twardzik DR, Vedvick TH;  
XX

DR WPI; 1997-192903/17.  
DR N-PSDB; AAT91521.  
XX  
PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are  
PT useful in vaccines for prevention or treatment of tuberculosis, also  
PT for diagnosis  
XX  
PS Example 3; Page 138-139; 168pp; English.  
XX  
CC A new immunogenic polypeptide has been developed comprising an  
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or  
CC its variant differing only in conservative substitutions and/or  
CC modifications). The present sequence represents a M. tuberculosis  
CC antigen, Tbh-9FL. The immunogenic protein, and fusion proteins  
CC containing one or more of the proteins or one of the proteins plus  
CC BSA-T-6, are useful in vaccines, preferably when formulated with a  
CC non-specific adjuvant, to induce an immune response against  
CC M. tuberculosis (for treatment or prevention).  
XX  
SQ Sequence 391 AA;  
Query Match 100.0%; Score 1949; DB 18; Length 391;  
Best Local Similarity 100.0%; Pred. No. 4.8e-143;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDGALPEINSARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGTLVGSWIG 60  
Db 1 MVDGALPEINSARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGTLVGSWIG 60  
QY 61 SSAGLWVAASPYVAMSVTAQAEELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
Db 61 SSAGLWVAASPYVAMSVTAQAEELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
QY 121 LIATNLIGONTPAIAVNEAEYGEWMAODAAAFGYAAATATATLLPFEAEEMTSAGG 180  
Db 121 LIATNLIGONTPAIAVNEAEYGEWMAODAAAFGYAAATATATLLPFEAEEMTSAGG 180  
QY 181 LLEQAAAVEASPTAAANQIMNNVPQALQOAPTOGTTPSSKLGIMKTVSPHRPISN 240  
Db 181 LLEQAAAVEASPTAAANQIMNNVPQALQOAPTOGTTPSSKLGIMKTVSPHRPISN 240  
QY 241 MVSMAHHMSTNSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300  
Db 241 MVSMAHHMSTNSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300  
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGQMLGGLPV 360  
Db 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPRPYVMPHSPAAG 391  
Db 361 GQMGARAGGGLSGVLRVPRPYVMPHSPAAG 391

RESULT 2  
AAM32381  
ID AAM32381 standard; Protein; 391 AA.  
XX  
AC AAM32381;  
XX  
DT 13-JAN-1998 (first entry)  
XX  
DE Mycobacterium tuberculosis antigen Tbh-9FL.  
XX  
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
XX skin testing; M. tuberculosis.  
OS Mycobacterium tuberculosis.  
XX  
XX MO9709429-A2.  
XX PN  
XX PD 13-MAR-1997.  
XX

PF 30-AUG-1996; 96MO-US14675.  
XX  
PR 12-JUL-1996; 96US-0680573.  
PR 01-SEP-1995; 95US-0523435.  
PR 22-SEP-1995; 95US-0532136.  
PR 22-MAR-1996; 96US-0620280.  
PR 05-JUN-1996; 96US-0658800.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;  
PI Twardzik DR, Vedvick TH;  
PI  
XX  
DR WPI; 1997-192904/17.  
DR N-PSDB; AAT91455.  
PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens  
PT - useful for diagnosis of M. tuberculosis infection  
XX  
PS Example 3; Page 150-152; 190pp; English.  
XX  
CC A new immunogenic polypeptide has been developed comprising an  
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or  
CC its variant differing only in conservative substitutions and/or  
CC modifications). The present sequence represents a M. tuberculosis  
CC antigen, Tbh-9FL. The immunogenic polypeptide can be used to diagnose  
CC M. tuberculosis infection by forming complexes with specific  
CC antibodies in the sample. Fragments of DNA encoding the immunogenic  
CC polypeptide can be used as diagnostic primers or probes and agents  
CC that bind to the antigen, especially monoclonal antibodies or  
CC equivalent polyclonal antibodies, are also used for diagnosis.  
XX  
SQ Sequence 391 AA;  
Query Match 100.0%; Score 1949; DB 18; Length 391;  
Best Local Similarity 100.0%; Pred. No. 4.8e-143;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDGALPEINSARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGTLVGSWIG 60  
Db 1 MVDGALPEINSARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGTLVGSWIG 60  
QY 61 SSAGLWVAASPYVAMSVTAQAEELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
Db 61 SSAGLWVAASPYVAMSVTAQAEELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
QY 121 LIATNLIGONTPAIAVNEAEYGEWMAODAAAFGYAAATATATLLPFEAEEMTSAGG 180  
Db 121 LIATNLIGONTPAIAVNEAEYGEWMAODAAAFGYAAATATATLLPFEAEEMTSAGG 180  
QY 181 LLEQAAAVEASPTAAANQIMNNVPQALQOAPTOGTTPSSKLGIMKTVSPHRPISN 240  
Db 181 LLEQAAAVEASPTAAANQIMNNVPQALQOAPTOGTTPSSKLGIMKTVSPHRPISN 240  
QY 241 MVSMAHHMSTNSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300  
Db 241 MVSMAHHMSTNSGVSMNTLSSMLKGFAPAAAQAQVOTPAQNGVRAMSSLGSSSG 300  
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGQMLGGLPV 360  
Db 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPRPYVMPHSPAAG 391  
Db 361 GQMGARAGGGLSGVLRVPRPYVMPHSPAAG 391

RESULT 3  
AAM81702  
ID AAM81702 standard; Protein; 391 AA.  
XX  
AC AAM81702;  
XX



DT 27-JAN-1999 (first entry)  
 XX  
 DE M. tuberculosis immunogenic polypeptide TBH-9FL.  
 XX  
 KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;  
 XX vaccine; pharmaceutical; infection; diagnosis.  
 OS  
 XX Mycobacterium tuberculosis.  
 XX  
 PN WO9816646-A2.  
 XX  
 PD 23-APR-1998.  
 XX  
 PF 07-OCT-1997; 97WO-US18293.  
 XX  
 PR 13-MAR-1997; 97US-0818112.  
 XX 11-OCT-1996; 96US-0730510.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;  
 PI Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS;  
 XX  
 DR WPI; 1998-261042/23.  
 DR N-PSDB; AAV64503.  
 XX  
 XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used  
 PT to develop products for the detection of M. tuberculosis infection  
 PT and for diagnosis, treatment and prevention of tuberculosis  
 XX  
 PS Example 3B; Page 128-129; 230pp; English.  
 XX  
 CC This sequence represents an immunogenic portion of a soluble  
 CC Mycobacterium tuberculosis (MT) antigen which can be used in a method  
 CC for inducing protective immunity against tuberculosis (TB). This  
 CC sequence can be formulated into vaccines and/or pharmaceutical  
 CC compositions for immunising against M. tuberculosis infection or may  
 CC be used for the diagnosis of tuberculosis.  
 CC  
 SQ Sequence 391 AA;  
 Query Match 100.0%; Score 1949; DB 19; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-143;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDFGALPEINSARNTAGPGSASLVAAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
 DB 1 MVDFGALPEINSARNTAGPGSASLVAAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
 QY 61 SSAGLMVAAASPYVAMSVTAQAELTAAQVRVAAAAYETAGLVPPVIAENRAELMI 120  
 DB 61 SSAGLMVAAASPYVAMSVTAQAELTAAQVRVAAAAYETAGLVPPVIAENRAELMI 120  
 QY 121 LIAATNLGONTPAIAVNEAEYGEEMAAQDAAMFGVAAATATATATLLPEEAPEMTSAGG 180  
 DB 121 LIAATNLGONTPAIAVNEAEYGEEMAAQDAAMFGVAAATATATATLLPEEAPEMTSAGG 180  
 QY 181 LIEQAAAVEASDTAAANQIMNNVPOALQOLAOPGTGTPSSKLGIMKTVPSPHRSPI 240  
 DB 181 LIEQAAAVEASDTAAANQIMNNVPOALQOLAOPGTGTPSSKLGIMKTVPSPHRSPI 240  
 QY 241 MYSMANHSMNTSGVSMNTLSSMLKGPAPAAAAYOTAQNGVRAVSSIGSSISG 300  
 DB 241 MYSMANHSMNTSGVSMNTLSSMLKGPAPAAAAYOTAQNGVRAVSSIGSSISG 300  
 QY 301 LGGGVAANTGRAASVGSLSVPQMAAANOAVTPAABALPLTSLTSAERGPQMLGLPLV 360  
 DB 301 LGGGVAANTGRAASVGSLSVPQMAAANOAVTPAABALPLTSLTSAERGPQMLGLPLV 360  
 QY 361 GQMGARAGGSLGVLRAVPPRPYVMPHSPAAG 391  
 DB 361 GQMGARAGGSLGVLRAVPPRPYVMPHSPAAG 391

RESULT 4  
 AAW64335  
 ID AAW64335 standard; Protein; 391 AA.  
 XX  
 AC AAW64335;  
 XX  
 DT 09-NOV-1998 (first entry)  
 XX  
 DE Mycobacterium tuberculosis antigen TBH-9FL.  
 XX  
 KW Tuberculosis; infection; diagnosis; antigen; TBH-9FL.  
 XX  
 OS Mycobacterium tuberculosis strain H37Rv.  
 XX  
 PN WO9816645-A2.  
 XX  
 PD 23-APR-1998.  
 XX  
 PF 07-OCT-1997; 97WO-US18214.  
 XX  
 PR 13-MAR-1997; 97US-0818111.  
 XX 11-OCT-1996; 96US-0729622.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;  
 PI Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS;  
 XX  
 DR WPI; 1998-251292/22.  
 DR N-PSDB; AAV44395.  
 XX  
 XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used  
 PT to develop products for the detection of M. tuberculosis infection  
 PT and diagnosis of tuberculosis  
 XX  
 PS Example 3; Page 133-135; 250pp; English.  
 XX  
 CC This polypeptide comprises Mycobacterium tuberculosis antigen  
 CC TBH-9FL. It is encoded by genomic DNA (see AAV44395) isolated from a  
 CC M. tuberculosis strain H37Rv genomic library using a probe from  
 CC clone TBH-9 (see AAV44371). The invention relates to compositions  
 CC and methods for diagnosing tuberculosis. It provides polypeptides  
 CC (see AAW64291-W64379) comprising an antigenic portion of a soluble  
 CC M. tuberculosis antigen, or an immunogenic portion of an M.  
 CC tuberculosis antigen, as well as DNA sequences encoding such  
 CC polypeptides, recombinant expression vectors and transformed or  
 CC transfected host cells. Also claimed are methods and diagnostic  
 CC kits for detecting M. tuberculosis infection in a patient using  
 CC these polypeptides, antibodies or oligonucleotide probes and  
 CC primers, for the diagnosis of tuberculosis.  
 CC  
 SQ Sequence 391 AA;  
 Query Match 100.0%; Score 1949; DB 19; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-143;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDFGALPEINSARNTAGPGSASLVAAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
 DB 1 MVDFGALPEINSARNTAGPGSASLVAAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
 QY 61 SSAGLMVAAASPYVAMSVTAQAELTAAQVRVAAAAYETAGLVPPVIAENRAELMI 120  
 DB 61 SSAGLMVAAASPYVAMSVTAQAELTAAQVRVAAAAYETAGLVPPVIAENRAELMI 120  
 QY 121 LIAATNLGONTPAIAVNEAEYGEEMAAQDAAMFGVAAATATATATLLPEEAPEMTSAGG 180  
 DB 121 LIAATNLGONTPAIAVNEAEYGEEMAAQDAAMFGVAAATATATATLLPEEAPEMTSAGG 180  
 QY 181 LIEQAAAVEASDTAAANQIMNNVPOALQOLAOPGTGTPSSKLGIMKTVPSPHRSPI 240  
 DB 181 LIEQAAAVEASDTAAANQIMNNVPOALQOLAOPGTGTPSSKLGIMKTVPSPHRSPI 240

```

QY 241 MVSANNHMTSGVSMNTLTSSMLKGFAPAPAAAQAQVOTAQNGVRAMSLGSSLSGSSG 300
DB 241 MVSANNHMTSGVSMNTLTSSMLKGFAPAPAAAQAQVOTAQNGVRAMSLGSSLSGSSG 300
QY 301 LGGGVAANLGRAASVGSLSVPQMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
DB 301 LGGGVAANLGRAASVGSLSVPQMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391
DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391

RESULT 5
AAV39132
ID AAV39132 standard; Protein; 391 AA.
XX AAV39132;
XX 05-NOV-1999 (first entry)
DE M. tuberculosis antigen Tbh-9FL amino acid sequence.
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KM immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test.
XX Mycobacterium tuberculosis.
XX WO9942076-A2.
XX 26-AUG-1999.
XX 17-FEB-1999; 99WO-US03268.
XX 05-MAY-1998; 98US-0072967.
XX 18-FEB-1998; 98US-0025197.
XX (CORI-) CORIXA CORP.
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX WPI: 1999-527409/44.
XX N-PSDB; AAZ19305.
XX New antigens from Mycobacterium tuberculosis useful in diagnostic
PT skin tests and protective or therapeutic vaccines or compositions
XX
PS Example 3; Page 123-124; 299pp; English.
XX
CC The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
CC are vaccines and fusion protein containing M. tuberculosis Ag's.
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
CC other polypeptides fragments, can be used in pharmaceutical compositions
CC or vaccines to generate a protective or therapeutic immune response to
CC M. tuberculosis and as reagents in skin tests for diagnosis of
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
CC by, T, B or natural killer cells and/or macrophages in
CC tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAV39083 to
CC AAV39225 are used in the exemplification of the present invention.
XX
SQ Sequence 391 AA;
Query Match 100.0%; Score 1949; DB 20; Length 391;
Best Local Similarity 100.0%; Pred. No. 4,8e-143;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MVDGALPPEINSAARMYAGPGSASLVAAQMWDSVADLFGSASAFQSVVWGLTVGSMIG 60
1 MVDGALPPEINSAARMYAGPGSASLVAAQMWDSVADLFGSASAFQSVVWGLTVGSMIG 60

```

```

QY 61 SSAGLWVAASPVVAMWSTAGQAEITTAQVRAAAAYETAYGLTVPPVIAENRAELMT 120
DB 61 SSAGLWVAASPVVAMWSTAGQAEITTAQVRAAAAYETAYGLTVPPVIAENRAELMT 120
QY 121 LIATNLIGONTPAIAVNEAEYGEEMAAQDAAMFGVAATATATATLLPFEAEEMTSAGG 180
DB 121 LIATNLIGONTPAIAVNEAEYGEEMAAQDAAMFGVAATATATATLLPFEAEEMTSAGG 180
QY 181 LLEQAAAVEASTTAANQIMNNVPAQLQQLAQPOTGTPSSSLGGLMKTIVSHRSPISN 240
DB 181 LLEQAAAVEASTTAANQIMNNVPAQLQQLAQPOTGTPSSSLGGLMKTIVSHRSPISN 240
QY 241 MVSANNHMTSGVSMNTLTSSMLKGFAPAPAAAQAQVOTAQNGVRAMSLGSSLSGSSG 300
DB 241 MVSANNHMTSGVSMNTLTSSMLKGFAPAPAAAQAQVOTAQNGVRAMSLGSSLSGSSG 300
QY 301 LGGGVAANLGRAASVGSLSVPQMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
DB 301 LGGGVAANLGRAASVGSLSVPQMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391
DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391

RESULT 6
AAV38989
ID AAV38989 standard; Protein; 391 AA.
XX AAV38989;
XX 05-NOV-1999 (first entry)
DE M. tuberculosis recombinant antigen protein Tbh-9FL.
XX Antigen; diagnosis; detection; infection; antibody; immunisation;
KM vaccine; immunity.
XX Mycobacterium tuberculosis.
XX WO9942118-A2.
XX 26-AUG-1999.
XX 17-FEB-1999; 99WO-US03265.
XX 05-MAY-1998; 98US-0072596.
XX 18-FEB-1998; 98US-0024753.
XX (CORI-) CORIXA CORP.
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX WPI: 1999-527416/44.
XX N-PSDB; AAZ19093.
XX New polypeptide comprising antigenic portions of M. tuberculosis
PT
PS Example 3; Page 168-169; 323pp; English.
XX
CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against
CC M. tuberculosis infection. The new detection methods are needed as
CC current vaccination strategies do not provide 100% immunity.
XX
SQ Sequence 391 AA;
Query Match 100.0%; Score 1949; DB 20; Length 391;
1 MVDGALPPEINSAARMYAGPGSASLVAAQMWDSVADLFGSASAFQSVVWGLTVGSMIG 60
1 MVDGALPPEINSAARMYAGPGSASLVAAQMWDSVADLFGSASAFQSVVWGLTVGSMIG 60

```

Best Local Similarity 100.0%; Pred. No. 4.8e-143;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQMDVSADLPSAASAFQSVVMGLTVGSWIG 60  
DB 1 MVDGALPPEINSARMYAGPGSASLVAAQMDVSADLPSAASAFQSVVMGLTVGSWIG 60  
QY 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
QY 121 LIATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
DB 121 LIATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
QY 181 LIEQAAAVEASDTAAANQLMNNVPQALQOLAOPTGTTSSKLGIMKTVPSPHRSPI SN 240  
DB 181 LIEQAAAVEASDTAAANQLMNNVPQALQOLAOPTGTTSSKLGIMKTVPSPHRSPI SN 240  
QY 241 MYSMANNHSMSTNSGVSMNTLTSSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 300  
DB 241 MYSMANNHSMSTNSGVSMNTLTSSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 300  
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360  
DB 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

## RESULT 7

AA04778  
ID AA04778 standard; Protein; 391 AA.

AC AA04778;

DT 06-JUL-1999 (first entry)

DE Mycobacterium species protein sequence 5R.

KM Secreted protein; Mycobacterium; primer; PCR; amplification; probe;  
XX hybridisation; detection; vaccine; immunisation; infection.

OS Mycobacterium sp.

PN WO9909186-A2.

PD 25-FEB-1999.

PF 14-AUG-1998; 98WO-FR01813.

PR 11-SEP-1997; 97FR-0011325.

PR 14-AUG-1997; 97FR-0010404.

XX (INSP ) INST PASTEUR.

PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;

DR N-PSDB; AAX34030.

XX MPI: 1999-181045/15.

PT Mycobacterial DNA vectors containing reporter constructs - for  
XX identifying coding or promoter sequences involved in  
XX infection-associated protein expression

PS Claim 32; Fig 5R; 309pp; French.

CC Sequences AA04742-Y05000 and AA07201-Y07204 represent secreted  
XX proteins from various Mycobacterium species microorganisms. The  
XX encoding nucleotide sequences can be used as primers and probes for

CC methods for detecting and identifying mycobacteria, especially belonging  
XX to the M. tuberculosis complex. The encoded proteins can be used in  
XX vaccines for immunisation against a bacterial or viral infection.

SQ Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 20; Length 391;

Best Local Similarity 100.0%; Pred. No. 4.8e-143;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQMDVSADLPSAASAFQSVVMGLTVGSWIG 60  
DB 1 MVDGALPPEINSARMYAGPGSASLVAAQMDVSADLPSAASAFQSVVMGLTVGSWIG 60  
QY 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
QY 121 LIATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
DB 121 LIATNLGONTPAIAVNEAEYEGEMMAQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
QY 181 LIEQAAAVEASDTAAANQLMNNVPQALQOLAOPTGTTSSKLGIMKTVPSPHRSPI SN 240  
DB 181 LIEQAAAVEASDTAAANQLMNNVPQALQOLAOPTGTTSSKLGIMKTVPSPHRSPI SN 240  
QY 241 MYSMANNHSMSTNSGVSMNTLTSSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 300  
DB 241 MYSMANNHSMSTNSGVSMNTLTSSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 300  
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360  
DB 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

## RESULT 8

AAU01888  
ID AAU01888 standard; Protein; 391 AA.

AC AAU01888;

DT 29-AUG-2001 (first entry)

DE M. tuberculosis antigen TbH9 (Mtb39A).

KM TbH9; Mtb39A; antigen; vaccine; tuberculosis; AIDS;  
XX acquired immunodeficiency disease.

OS Mycobacterium tuberculosis.

PN MO200124820-A1.

PD 12-APR-2001.

PF 10-OCT-2000; 2000WO-US28095.

PR 07-OCT-1999; 99US-0158338.

PR 07-OCT-1999; 99US-0158425.

XX (CORI-) CORIXA CORP.

PI Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;

DR MPI: 2001-290576/30.

DR N-PSDB; AAS03779.

XX Vaccinating against Mycobacteria infections in mammals using fusion  
XX proteins comprising combinations of heterologous antigens -



```

XX  MO200198460-A2.
XX
XX  27-DEC-2001.
XX
XX  20-JUN-2001; 2001WO-US19959.
XX
XX  20-JUN-2000; 2000US-0597796.
XX  01-FEB-2001; 2001US-265737P.
XX  (CORI-) CORIXA CORP.
XX
XX  Skelky Y, Reed S, Alderson M;
XX  WPI: 2002-147798/19.
XX  N-PSDB; AAD28341.
XX
XX  Composition comprising MTB39 antigen and MTB32A antigen from
XX  Mycobacterium species, useful for eliciting immune response in a
XX  subject
XX
XX  Claim 83; Page 102-103; 136pp; English.
XX
XX  The present invention relates to fusion proteins containing at least
XX  two Mycobacterium species antigens, nucleotides encoding them and
XX  compositions comprising such fusion proteins. The present invention
XX  particularly relates to nucleic acids encoding fusion proteins that
XX  include two or more individual M. tuberculosis antigens which increase
XX  the serological sensitivity of sera from individuals infected with
XX  tuberculosis and methods for their use in diagnosis, prevention and
XX  treatment of tuberculosis infection. Sequences of the invention are
XX  useful for eliciting an immune response in a mammal, e.g., human,
XX  immunised with BCG. They are useful in the diagnosis, treatment and
XX  prevention of Mycobacterium infection. The fusion proteins and the
XX  polynucleotides are useful as diagnostic tools in patients infected
XX  with Mycobacterium, in vitro and in vivo assays for detecting humoral
XX  antibodies or cell-mediated immunity against M. tuberculosis, for the
XX  diagnosis of an infection or monitoring of disease progression, as
XX  immunogens to generate or elicit a protective immune response in a
XX  patient and for raising anti-M. tuberculosis antibodies in a non-human
XX  animal. Sequences of the invention are also used as vaccines. MTB32A
XX  fusion proteins of the invention are useful as in vivo diagnostic agents
XX  for intradermal skin test. The present sequence is Mycobacterium species
XX  MTB39 (TbH9) protein.
XX
XX  Sequence 391 AA;
XX
XX  Query Match 100.0%; Score 1949; DB 23; Length 391;
XX  Best Local Similarity 100.0%; Pred. No. 4,8e-143;
XX  Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB  301 LGGVAAANLGRAASVGLSVPOMAAANQAVTPAARALPLTSLTSAERGPGOMLGGLPV 360
QY  361 GOMGARAGGGLSGVLRVPRPYVMPHSPAG 391
DB  361 GOMGARAGGGLSGVLRVPRPYVMPHSPAG 391

RESULT 11
AA32070
ID  AA32070 standard; Protein; 596 AA.
XX
XX  AA32070;
XX
XX  17-JAN-2000 (first entry)
XX
XX  Mycobacterium tuberculosis antigen fusion protein Mtb59f.
XX  Tuberculosis; antigen; fusion protein; Mtb59f; TbH9; Ra35;
XX  diagnosis; therapy; vaccine; immunogen.
XX
XX  Mycobacterium tuberculosis.
XX
XX  Key Location/Qualifiers
XX  FH Peptide 1..8
XX  FT /note= "Met/His tag"
XX  FT Protein 9..140
XX  FT /note= "Ra12"
XX  FT Protein 143..596
XX  FT /note= "TbH9"
XX
XX  WO951748-A2.
XX
XX  14-OCT-1999.
XX
XX  07-APR-1999; 99WO-US07717.
XX
XX  07-APR-1998; 98US-0056556.
XX  30-DEC-1998; 98US-0223040.
XX  (CORI-) CORIXA CORP.
XX
XX  Skelky YAW, Alderson M, Campos-Neto A;
XX  WPI: 1999-601610/51.
XX  N-PSDB; AAZ20205.
XX
XX  New fusion proteins useful for diagnosis, prevention and treatment of
XX  tuberculosis -
XX
XX  Claim 1; Fig 12A-B; 83pp; English.
XX
XX  This sequence represents a recombinant Mycobacterium tuberculosis
XX  bi-antigen fusion protein, termed Mtb59f, composed of the antigens
XX  TbH9 and Ra35. The fusion protein is expressed in host cells
XX  using a vector carrying a polynucleotide (see AAZ20205) comprising
XX  the coding sequences for the 2 antigens. The invention provides
XX  fusion proteins (see AA32059-71) containing at least 2 M.
XX  tuberculosis antigens. The new fusion proteins and polynucleotides
XX  encoding them are useful as vaccines for preventing tuberculosis
XX  (claimed), for diagnosis (via in vitro assays or intradermal skin
XX  tests for detection of anti-M. tuberculosis antibodies), monitoring
XX  of disease progression, and treatment of tuberculosis. They are
XX  more effective immunogens than mixtures of the individual protein
XX  components.
XX
XX  Sequence 596 AA;
XX
XX  Query Match 100.0%; Score 1949; DB 20; Length 596;
XX  Best Local Similarity 100.0%; Pred. No. 8.1e-143;
XX  Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 9 MVDGALPEINSAARMYAGPGSASLVAAQMDSVASDLFSAASAFQSVMGLTVGSWIG 68  
 QY 61 SSAGLWVAASPYVAMMSVTAGAEITTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 120  
 Db 69 SSAGLWVAASPYVAMMSVTAGAEITTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 128  
 QY 121 LIATNLLGQWTPAIAVNEAEYGEWMAODAAAFGYAAATATATATLLPFEAPEMTSAGG 180  
 Db 129 LIATNLLGQWTPAIAVNEAEYGEWMAODAAAFGYAAATATATATLLPFEAPEMTSAGG 188  
 QY 181 LLEQAAAVEEASDTPAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPSPHSPISN 240  
 Db 189 LLEQAAAVEEASDTPAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPSPHSPISN 248  
 QY 241 MVSANNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSSG 300  
 Db 249 MVSANNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSSG 308  
 QY 301 LGGGVAANLGRAASVGSLSVQMAAANQAVTPARALPLTSLTSAERPGQMLGGLPV 360  
 Db 309 LGGGVAANLGRAASVGSLSVQMAAANQAVTPARALPLTSLTSAERPGQMLGGLPV 368  
 QY 361 GQMGARAGGSLGVLVPPRPVMPHSPAG 391  
 Db 369 GQMGARAGGSLGVLVPPRPVMPHSPAG 399

## RESULT 12

AAE29710 standard; Protein; 596 AA.

AAE29710;

27-JAN-2003 (first entry)

Mycobacterium sp. MTB59F fusion protein.

Vaccine; immunity; diagnostic agent; gene therapy; TBH9; antigen;

Ra35; MTB59F; fusion protein.

Chimeric - Mycobacterium sp.

Chimeric - Mycobacterium tuberculosis.

W0200272792-A2.

19-SEP-2002.

13-MAR-2002; 2002WO-US08223.

13-MAR-2001; 2001US-275837P.

(CORI-) CORIXA CORP.

Skeiky Y, Brannon M, Guderian J;

WPI; 2002-759844/82.

N-PSDB; AAD47086.

New recombinant nucleic acid molecule comprising a Leishmania TSA.

LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective

immunity against pathogenic microorganisms e.g. Leishmania and

Mycobacterium tuberculosis

Disclousure; Page 98-99; 155pp; English.

The invention relates to a recombinant nucleic acid molecule encoding a

fusion polypeptide. The recombinant nucleic acid comprises a heterologous

polynucleotide sequence encoding an antigen or an antigenic fragment from

Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a

polypeptide or its fragment. The Leishmania polynucleotide is selected

from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention

are used in methods for eliciting immune response in mammals. They are

useful as vaccines to elicit protective immunity against pathogenic

CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
 CC polypeptides are used for enhancing the expression of polynucleotides,  
 CC as in vivo diagnostic agents and for raising antibodies in a non-human  
 CC animal. The invention is used in gene therapy. The present sequence is  
 CC MTB59F fusion protein. This fusion protein comprises Ra35 protein from  
 CC Mycobacterium tuberculosis and TBH9 protein from Mycobacterium sp.

SQ Sequence 596 AA;

Query Match 100.0%; Score 1949; DB 23; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 8, 1e-143;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARMYAGPGSASLVAAQMDSVASDLFSAASAFQSVMGLTVGSWIG 60  
 Db 9 MVDGALPEINSAARMYAGPGSASLVAAQMDSVASDLFSAASAFQSVMGLTVGSWIG 68  
 QY 61 SSAGLWVAASPYVAMMSVTAGAEITTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 120  
 Db 69 SSAGLWVAASPYVAMMSVTAGAEITTAQVRAVAAAAYETAYGLTVPPVIAENRAELMI 128  
 QY 121 LIATNLLGQWTPAIAVNEAEYGEWMAODAAAFGYAAATATATATLLPFEAPEMTSAGG 180  
 Db 129 LIATNLLGQWTPAIAVNEAEYGEWMAODAAAFGYAAATATATATLLPFEAPEMTSAGG 188  
 QY 181 LLEQAAAVEEASDTPAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPSPHSPISN 240  
 Db 189 LLEQAAAVEEASDTPAANQIMNNVPOALQOLAQPTGTTSSKLGIMKTVPSPHSPISN 248  
 QY 241 MVSANNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSSG 300  
 Db 249 MVSANNHMSMTNSGVSMNTTLLSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSSG 308  
 QY 301 LGGGVAANLGRAASVGSLSVQMAAANQAVTPARALPLTSLTSAERPGQMLGGLPV 360  
 Db 309 LGGGVAANLGRAASVGSLSVQMAAANQAVTPARALPLTSLTSAERPGQMLGGLPV 368  
 QY 361 GQMGARAGGSLGVLVPPRPVMPHSPAG 391  
 Db 369 GQMGARAGGSLGVLVPPRPVMPHSPAG 399

## RESULT 13

AAE17574 standard; Protein; 596 AA.

AAE17574;

22-APR-2002 (first entry)

Mycobacterium species MTB59F fusion protein.

Fusion protein; antigen; serological sensitivity; immune response;

tuberculosis; infection; vaccine; MTB59F; TBH9-Ra35 protein.

Mycobacterium sp.

W0200198460-A2.

20-JUN-2001; 2001WO-US19959.

20-JUN-2000; 2000US-0597796.

01-FEB-2001; 2001US-265737P.

(CORI-) CORIXA CORP.

Skeiky Y, Reed S, Alderson M;

WPI; 2002-147798/19.

N-PSDB; AAD28344.

PT Composition comprising MTB39 antigen and MTB32A antigen from  
 PT Mycobacterium species, useful for eliciting immune response in a  
 PT subject  
 XX  
 PS Claim 5; Page 114-115; 136pp; English.  
 CC The present invention relates to fusion proteins containing at least  
 CC two Mycobacterium species antigens, nucleotides encoding them and  
 CC compositions comprising such fusion proteins. The present invention  
 CC particularly relates to nucleic acids encoding fusion proteins that  
 CC include two or more individual M. tuberculosis antigens which increase  
 CC the serological sensitivity of sera from individuals infected with  
 CC tuberculosis and methods for their use in diagnosis, prevention and  
 CC treatment of tuberculosis infection. Sequences of the invention are  
 CC useful for eliciting an immune response in a mammal, e.g., human,  
 CC immunised with BCG. They are useful in the diagnosis, treatment and  
 CC prevention of Mycobacterium infection. The fusion proteins and the  
 CC polynucleotides are useful as diagnostic tools in patients infected  
 CC with Mycobacterium, in vitro and in vivo assays for detecting humoral  
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
 CC diagnosis of an infection or monitoring of disease progression, as  
 CC immunogens to generate or elicit a protective immune response in a  
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
 CC animal. Sequences of the invention are also used as vaccines. MTB32A  
 CC fusion proteins of the invention are useful as in vivo diagnostic agents  
 CC for intradermal skin test. The present sequence is Mycobacterium species  
 CC MTB59F (TbH9-Ra35) fusion protein.  
 XX  
 SQ Sequence 596 AA;  
 Query Match 100.0%; Score 1949; DB 23; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-143;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MWDGALPPEINSARMYAGPGSASLVAAQMDVSADLFSASAFQSVVMGLTVGSWIG 60  
 DB 9 MWDGALPPEINSARMYAGPGSASLVAAQMDVSADLFSASAFQSVVMGLTVGSWIG 68  
 QY 61 SSAGLMVAAASPYVAMMSVTAGQAEITTAQVRAAAAYETAGLTVPPVIAENRAELMI 120  
 DB 69 SSAGLMVAAASPYVAMMSVTAGQAEITTAQVRAAAAYETAGLTVPPVIAENRAELMI 128  
 QY 121 LIATNLGONTPAIAVNEAEYEGEMMAODAAAMFGVAAATATATATLLPPEBAPEMTSAG 180  
 DB 129 LIATNLGONTPAIAVNEAEYEGEMMAODAAAMFGVAAATATATATLLPPEBAPEMTSAG 188  
 QY 181 LLEQAAAVEASDTAAANOLMNNVPQALQALOPTGTPSSKLGIMKTVPSPHRSPISN 240  
 DB 189 LLEQAAAVEASDTAAANOLMNNVPQALQALOPTGTPSSKLGIMKTVPSPHRSPISN 248  
 QY 241 MWSMANNHSMNTSGVSMNTLTLSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 300  
 DB 249 MWSMANNHSMNTSGVSMNTLTLSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 308  
 QY 301 IGGGVAAANGRAASVGSLSVPOAMAAANQAVTPARALPTLSLSAERGGOMLGGLPV 360  
 DB 309 IGGGVAAANGRAASVGSLSVPOAMAAANQAVTPARALPTLSLSAERGGOMLGGLPV 368  
 QY 361 GQWGARAGGGLGCLRVPRPYVMPHSPAAG 391  
 DB 369 GQWGARAGGGLGCLRVPRPYVMPHSPAAG 399  
 RESULT 14  
 AAU74599  
 ID AAU74599 standard; Protein; 599 AA.  
 XX AAU74599;  
 AC  
 XX 08-MAY-2002 (first entry)  
 DT  
 XX Antigenic fusion protein Tb59-Ra35 (Meb59F).  
 DE  
 XX

KW Fusion protein; tuberculosis; Mycobacterium tuberculosis;  
 KW tuberculosis; immunogen; vaccine; Tb59-Ra35; Meb59F.  
 XX  
 OS Chimeric - Mycobacterium tuberculosis.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 597  
 FT /label= OTHER  
 FT /note= "OTHER= Xaa. Xaa= In frame stop codon"  
 XX  
 PN US2002009459-A1.  
 PD  
 PD 24-JAN-2002.  
 XX  
 PF 07-APR-1999; 99US-0287849.  
 XX  
 PR 13-MAR-1997; 97US-0818112.  
 PR 01-OCT-1997; 97US-0942578.  
 PR 18-FEB-1998; 98US-0025197.  
 PR 07-APR-1998; 98US-0056556.  
 PR 30-DEC-1998; 98US-0223040.  
 XX  
 PA (REED/) REED S G.  
 PA (SKEI/) SKEIKY Y A.  
 PA (DILL/) DILLON D C.  
 PA (ALDE/) ALDERSON M.  
 PA (CAMP/) CAMPOS-NETO A.  
 XX  
 PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
 XX  
 DR WPI; 2002-171134/22.  
 DR N-PSDB; ABL14139.  
 XX  
 PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for  
 PT diagnosing, treating or preventing M. tuberculosis infection,  
 PT particularly as vaccine for treating or preventing tuberculosis  
 XX  
 PS Claim 1; Fig 12; 62pp; English.  
 XX  
 CC The invention relates to a purified polypeptide which induces an immune  
 CC response of Mycobacterium tuberculosis. Polypeptides of the invention are  
 CC useful for diagnosing, treating or preventing M. tuberculosis infection,  
 CC particularly tuberculosis infection. In particular, the polypeptides are  
 CC useful as a vaccine formulation with an adjuvant to afford long-term  
 CC protection in animals against the development of tuberculosis. The  
 CC protein coding sequence may be used to encode a protein product for use  
 CC as an immunogen to induce and/or enhance an immune response to M.  
 CC tuberculosis. This sequence represents an M. tuberculosis fusion protein  
 CC of the invention.  
 XX  
 SQ Sequence 599 AA;  
 Query Match 100.0%; Score 1949; DB 23; Length 599;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-143;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MWDGALPPEINSARMYAGPGSASLVAAQMDVSADLFSASAFQSVVMGLTVGSWIG 60  
 DB 9 MWDGALPPEINSARMYAGPGSASLVAAQMDVSADLFSASAFQSVVMGLTVGSWIG 68  
 QY 61 SSAGLMVAAASPYVAMMSVTAGQAEITTAQVRAAAAYETAGLTVPPVIAENRAELMI 120  
 DB 69 SSAGLMVAAASPYVAMMSVTAGQAEITTAQVRAAAAYETAGLTVPPVIAENRAELMI 128  
 QY 121 LIATNLGONTPAIAVNEAEYEGEMMAODAAAMFGVAAATATATATLLPPEBAPEMTSAG 180  
 DB 129 LIATNLGONTPAIAVNEAEYEGEMMAODAAAMFGVAAATATATATLLPPEBAPEMTSAG 188  
 QY 181 LLEQAAAVEASDTAAANOLMNNVPQALQALOPTGTPSSKLGIMKTVPSPHRSPISN 240  
 DB 189 LLEQAAAVEASDTAAANOLMNNVPQALQALOPTGTPSSKLGIMKTVPSPHRSPISN 248  
 QY 241 MWSMANNHSMNTSGVSMNTLTLSMLKGFAPAAAQAQVOTAAQNGVRAMSLGSSIGSSG 300

Db 249 MWSMANNHMSMTNSGVSMNTLTSSMLKGFAPAAAQAVTAQNGVRAMSSLGSSSG 308  
QY 301 IGGGVAANLGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
Db 309 IGGGVAANLGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 368  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAG 391  
Db 369 GQMGARAGGGLSGVLRVPPRPVMPHSPAG 399

RESULT 15  
AAI32068  
ID AAI32068 standard; Protein; 600 AA.  
XX AAI32068;  
AC AAI32068;  
XX 17-JAN-2000 (first entry)  
DT 17-JAN-2000 (first entry)  
DE Mycobacterium tuberculosis antigen fusion protein Mtb61f.  
XX Tuberculosis; antigen; fusion protein; Mtb61f; TBH9; DPV; MTI;  
KM diagnosis; therapy; vaccine; immunogen.  
XX Mycobacterium tuberculosis.  
OS Mycobacterium tuberculosis.  
XX MO9951748-A2.  
XX PN 14-OCT-1999.  
XX PD 14-OCT-1999.  
XX PF 07-APR-1999; 99WO-US07717.  
XX PR 07-APR-1998; 98US-0056556.  
XX PR 30-DEC-1998; 98US-0223040.  
XX (CORI-) CORIXA CORP.  
PA Skeiky YAM, Alderson M, Campos-Neto A;  
PI WPI; 1999-601610/51.  
XX DR N-PSDB; AAZ20203.  
XX New fusion proteins useful for diagnosis, prevention and treatment of  
PT tuberculosis -  
XX Claim 1; Fig 10A-B; 83pp; English.  
PS This sequence represents a recombinant Mycobacterium tuberculosis  
XX tri-antigen fusion protein, termed Mtb61f, composed of the antigens  
CC TBH9, DPV and MTI. The fusion protein is expressed in host cells  
CC using a vector carrying a polynucleotide (see AAZ20203) comprising  
CC the coding sequences for the 3 antigens. The invention provides  
CC fusion proteins (see AAI32059-71) containing at least 2 M.  
CC tuberculosis antigens. The new fusion proteins and polynucleotides  
CC encoding them are useful as vaccines for preventing tuberculosis  
CC (claimed), for diagnosis (via in vitro assays or intradermal skin  
CC tests for detection of anti-M. tuberculosis antibodies), monitoring  
CC of disease progression, and treatment of tuberculosis. They are  
CC more effective immunogens than mixtures of the individual protein  
XX components.  
XX  
SQ Sequence 600 AA;

Query Match 100.0%; Score 1949; DB 20; Length 600;  
Best Local Similarity 100.0%; Pred. No. 8.2e-143;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMTAGPSASLVAQAQWDSVASDLPFAASAFQSVVWGLTVGSMIG 60  
Db 9 MVDGALPPEINSARMTAGPSASLVAQAQWDSVASDLPFAASAFQSVVWGLTVGSMIG 68  
QY 61 SSAGLWVAASPYVAMSVTAQAELTAQVRVAAAAYETAYGLTPPPVIAENRAELMT 120

Db 69 SSAGLWVAASPYVAMSVTAQAELTAQVRVAAAAYETAYGLTPPPVIAENRAELMT 128  
QY 121 LIATNLLGQVTPAIAVNEAEYGEWMAODAAAFGVAATATATATLLPFEAEDEMTSAGG 180  
Db 129 LIATNLLGQVTPAIAVNEAEYGEWMAODAAAFGVAATATATATLLPFEAEDEMTSAGG 188  
QY 181 LLEQAAAVEEASDPTAAANQIMNNVPOALQOLAQPTGCTPSSKLGLMKTVSFHRSPISN 240  
Db 189 LLEQAAAVEEASDPTAAANQIMNNVPOALQOLAQPTGCTPSSKLGLMKTVSFHRSPISN 248  
QY 241 MWSMANNHMSMTNSGVSMNTLTSSMLKGFAPAAAQAVTAQNGVRAMSSLGSSSG 300  
Db 249 MWSMANNHMSMTNSGVSMNTLTSSMLKGFAPAAAQAVTAQNGVRAMSSLGSSSG 308  
QY 301 IGGGVAANLGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
Db 309 IGGGVAANLGRAASVGSLSVPOMAAANAQAVTPAARALPLTSLTSAERPGQMLGGLPV 368  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAG 391  
Db 369 GQMGARAGGGLSGVLRVPPRPVMPHSPAG 399

Search completed: November 21, 2003, 17:04:07  
Job time : 46 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:03:15 ; Search time 21 Seconds  
(without alignments)  
787.788 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949  
Sequence: 1 MVDFCALPEPEINSARMYAGP.....SGVLKVPPEPYMPHPAAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1949	100.0	391	3	US-08-818-112-107 Sequence 107, App
2	1949	100.0	391	4	US-08-818-111-102 Sequence 102, App
3	1949	100.0	391	4	US-09-056-556-107 Sequence 107, App
4	1949	100.0	391	4	US-09-072-556-102 Sequence 102, App
5	1944	99.7	729	4	US-09-223-040-2 Sequence 2, App1
6	1652.5	84.8	396	3	US-08-818-112-111 Sequence 111, App
7	1652.5	84.8	396	4	US-08-818-111-106 Sequence 106, App
8	1652.5	84.8	396	4	US-09-056-556-111 Sequence 111, App
9	1652.5	84.8	396	3	US-09-072-556-106 Sequence 106, App
10	1486.5	76.3	359	4	US-08-818-112-109 Sequence 109, App
11	1486.5	76.3	359	4	US-08-818-111-104 Sequence 104, App
12	1486.5	76.3	359	4	US-09-056-556-109 Sequence 109, App
13	1486.5	76.3	359	4	US-09-072-556-104 Sequence 104, App
14	1187	60.9	263	3	US-08-818-112-91 Sequence 91, App1
15	1187	60.9	263	4	US-08-818-111-92 Sequence 92, App1
16	1187	60.9	263	4	US-09-056-556-91 Sequence 91, App1
17	1187	60.9	263	4	US-09-072-556-92 Sequence 92, App1
18	766.5	39.3	400	4	US-09-073-009-126 Sequence 126, App
19	603	30.9	423	4	US-09-073-009-142 Sequence 142, App
20	424.5	21.8	943	4	US-09-477-135A-131 Sequence 131, App
21	424	21.8	141	4	US-09-073-009-15 Sequence 15, App1
22	381.5	19.6	204	4	US-08-311-731A-57 Sequence 57, App1
23	377.5	19.4	208	4	US-08-311-731A-208 Sequence 208, App
24	314	16.1	368	3	US-08-818-112-114 Sequence 114, App
25	314	16.1	368	4	US-08-818-111-109 Sequence 109, App
26	314	16.1	368	4	US-09-056-556-114 Sequence 114, App
27	314	16.1	368	4	US-09-072-556-109 Sequence 109, App

28	261	13.4	1271	1	US-08-095-734-2	Sequence 2, App1
29	261	13.4	1271	2	US-08-444-623-2	Sequence 2, App1
30	261	13.4	1271	3	US-08-471-869-2	Sequence 2, App1
31	261	13.4	1271	4	US-09-342-563-2	Sequence 2, App1
32	261	13.4	1271	5	PCT-US94-08267-2	Sequence 2, App1
33	215	11.0	352	4	US-09-073-009-14	Sequence 14, App1
34	186.5	9.6	943	4	US-09-056-556-204	Sequence 204, App
35	186.5	9.6	943	4	US-09-072-556-199	Sequence 199, App
36	145	7.4	800	4	US-09-252-991A-20437	Sequence 20437, A
37	143.5	7.4	228	4	US-09-477-135A-128	Sequence 128, App
38	141	7.2	738	3	US-08-864-038A-3	Sequence 3, App1
39	139.5	7.2	792	2	US-08-678-039A-40	Sequence 40, App1
40	131.5	6.7	826	4	US-08-894-998A-47	Sequence 47, App1
41	129.5	6.6	1186	2	US-08-861-464-8	Sequence 8, App1
42	129.5	6.6	1186	2	US-08-396-001-8	Sequence 8, App1
43	129.5	6.6	1186	3	US-09-323-433A-8	Sequence 8, App1
44	128	6.6	731	4	US-09-340-736E-1	Sequence 1, App1
45	126	6.5	731	2	US-08-911-364-1	Sequence 1, App1

## ALIGNMENTS

RESULT 1  
US-08-818-112-107

; Sequence 107, Application US/08818112  
; Patent No. 6290969

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio

APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.

APPLICANT: Twadzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:  
ADDRESSER: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle

STATE: Washington  
COUNTRY: USA

ZIP: 98104-7092  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,112

FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.

REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C6

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids

TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
US-08-818-112-107

Query Match 100.0%; Score 1949; DB 3; Length 391;  
Best Local Similarity 100.0%; Pred. No. 2,7e-154;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARVYAGPGSASLVAAQMDVSASDLFSAASAFQSVVMGLTVGSMIG 60  
DB 1 MVDGALPEINSARVYAGPGSASLVAAQMDVSASDLFSAASAFQSVVMGLTVGSMIG 60  
QY 61 SSAGLMVAASAPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
DB 61 SSAGLMVAASAPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
QY 121 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPFEBAPEMISAGG 180  
DB 121 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPFEBAPEMISAGG 180  
QY 181 LLEQAAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPIIN 240  
DB 181 LLEQAAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPIIN 240  
QY 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSIGSSG 300  
DB 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSIGSSG 300  
QY 301 LGGGVANILGRAASVGSLSVPOAMAAANQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
DB 301 LGGGVANILGRAASVGSLSVPOAMAAANQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 391

## RESULT 2

US-08-818-111-102  
; Sequence 102, Application US/08818111  
; Patent No. 6338852

## GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,111  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-818-111-102

Query Match 100.0%; Score 1949; DB 4; Length 391;  
Best Local Similarity 100.0%; Pred No. 2.7e-154;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARVYAGPGSASLVAAQMDVSASDLFSAASAFQSVVMGLTVGSMIG 60  
DB 1 MVDGALPEINSARVYAGPGSASLVAAQMDVSASDLFSAASAFQSVVMGLTVGSMIG 60  
QY 61 SSAGLMVAASAPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
DB 61 SSAGLMVAASAPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
QY 121 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPFEBAPEMISAGG 180  
DB 121 LATNLGONTPAIAVNEAEYGEEMWADAAAFGYAAATATATATLLPFEBAPEMISAGG 180  
QY 181 LLEQAAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPIIN 240  
DB 181 LLEQAAAVEASDTAAANQLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPIIN 240  
QY 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSIGSSG 300  
DB 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSIGSSG 300  
QY 301 LGGGVANILGRAASVGSLSVPOAMAAANQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
DB 301 LGGGVANILGRAASVGSLSVPOAMAAANQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAG 391

## RESULT 3

US-09-056-556-107  
; Sequence 107, Application US/09056556  
; Patent No. 6350456

## GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-09-056-556-107

TREATM

US-09-056-556-107

Query Match 100.0%; Score 1949; DB 4; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-154;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 61 SSAGLMVAAASPYVAMSVTAGQAEELTAQVRAAAVETAYGLTVPPVIAENRAELMI 120  
 61 SSAGLMVAAASPYVAMSVTAGQAEELTAQVRAAAVETAYGLTVPPVIAENRAELMI 120  
 121 LIATNLGONTPALAVNEAEYGEWMAODAAAFGYAAATATATATLLPPEEAPEMTSAAG 180  
 121 LIATNLGONTPALAVNEAEYGEWMAODAAAFGYAAATATATATLLPPEEAPEMTSAAG 180  
 121 LIATNLGONTPALAVNEAEYGEWMAODAAAFGYAAATATATATLLPPEEAPEMTSAAG 180  
 181 LLEQAAVEASDPTAAANQOLMNNVPOALQOLAOPTQGTTPSSKLGMLKTVSPHRSPISN 240  
 181 LLEQAAVEASDPTAAANQOLMNNVPOALQOLAOPTQGTTPSSKLGMLKTVSPHRSPISN 240  
 241 MWSMANNHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSLSGSG 300  
 241 MWSMANNHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSLSGSG 300  
 301 LGGVAAANLGRPAASVGSISVPOAMAAANQAVTPARALPLTSLTSAERGPOMIGLPIV 360  
 301 LGGVAAANLGRPAASVGSISVPOAMAAANQAVTPARALPLTSLTSAERGPOMIGLPIV 360  
 361 GOMGARAGGSLGVLRVPPRPVMPHSPAG 391  
 361 GOMGARAGGSLGVLRVPPRPVMPHSPAG 391

RESULT 4

US-09-072-596-102

Sequence 102; Application US/09072596  
 Patent No. 6458356

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Campos-Neto, Antonia  
 APPLICANT: Houghton, Raymond  
 APPLICANT: Vedvick, Thomas S.  
 APPLICANT: Twardzik, Daniel R.  
 APPLICANT: Lodes, Michael J.  
 APPLICANT: Hendrickson, Ronald C.  
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
 NUMBER OF SEQUENCES: 350  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED AND BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,596  
 FILING DATE: 05-MAY-1998

CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.417C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 391 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-09-072-596-102

Query Match 100.0%; Score 1949; DB 4; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-154;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 61 SSAGLMVAAASPYVAMSVTAGQAEELTAQVRAAAVETAYGLTVPPVIAENRAELMI 120  
 61 SSAGLMVAAASPYVAMSVTAGQAEELTAQVRAAAVETAYGLTVPPVIAENRAELMI 120  
 121 LIATNLGONTPALAVNEAEYGEWMAODAAAFGYAAATATATATLLPPEEAPEMTSAAG 180  
 121 LIATNLGONTPALAVNEAEYGEWMAODAAAFGYAAATATATATLLPPEEAPEMTSAAG 180  
 121 LIATNLGONTPALAVNEAEYGEWMAODAAAFGYAAATATATATLLPPEEAPEMTSAAG 180  
 181 LLEQAAVEASDPTAAANQOLMNNVPOALQOLAOPTQGTTPSSKLGMLKTVSPHRSPISN 240  
 181 LLEQAAVEASDPTAAANQOLMNNVPOALQOLAOPTQGTTPSSKLGMLKTVSPHRSPISN 240  
 241 MWSMANNHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSLSGSG 300  
 241 MWSMANNHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLGSSLSGSG 300  
 301 LGGVAAANLGRPAASVGSISVPOAMAAANQAVTPARALPLTSLTSAERGPOMIGLPIV 360  
 301 LGGVAAANLGRPAASVGSISVPOAMAAANQAVTPARALPLTSLTSAERGPOMIGLPIV 360  
 361 GOMGARAGGSLGVLRVPPRPVMPHSPAG 391  
 361 GOMGARAGGSLGVLRVPPRPVMPHSPAG 391

RESULT 5

US-09-223-040-2

Sequence 2; Application US/09223040  
 Patent No. 6544522

GENERAL INFORMATION:  
 APPLICANT: Skeiky, Yasir  
 APPLICANT: Alderson, Mark  
 APPLICANT: Campos-Neto, Antonio  
 APPLICANT: Corixa Corporation  
 TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 TITLE OF INVENTION: and Their Uses  
 FILE REFERENCE: 014058-009010US  
 CURRENT APPLICATION NUMBER: US/09/223,040  
 CURRENT FILING DATE: 1998-12-30  
 NUMBER OF SEQ ID NOS: 10  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 2  
 LENGTH: 729  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: tri-fusion

Query Match 99.7%; Score 1944; DB 4; Length 729;  
 Best Local Similarity 99.7%; Pred. No. 1.7e-153;  
 Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 60  
 142 MVDGALPPEINSAMVAGPGSASLVAAAQWMDVSASDLFSAASAFQSVVMGLTVGSMWG 201

QY 61 SSAGLWVAASPVVAMSVTGAQELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
| | | | |  
Db 202 SSAGLWVAASPVVAMSVTGAQELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 261  
| | | | |  
QY 121 LATNLGONTPAIAVNEAEYEGEMWADAAMFGVAATAATATATLLPFEAEEMTSAGG 180  
| | | | |  
Db 262 LATNLGONTPAIAVNEAEYEGEMWADAAMFGVAATAATATATLLPFEAEEMTSAGG 321  
| | | | |  
QY 181 LLEQAAVVEASDTAAANQMLNNVPALQOLAPTOGTTSSKLGLMKTVSHRSPISN 240  
| | | | |  
Db 322 LLEQAAVVEASDTAAANQMLNNVPALQOLAPTOGTTSSKLGLMKTVSHRSPISN 381  
| | | | |  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVTAQNGVRAMSSLGSSLSG 300  
| | | | |  
Db 382 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVTAQNGVRAMSSLGSSLSG 441  
| | | | |  
QY 301 LGGGVAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
| | | | |  
Db 442 LGGGVAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 501  
| | | | |  
QY 361 GQMGARAGGSLGVLPVPPPYMPHSPAG 391  
| | | | |  
Db 502 GQMGARAGGSLGVLPVPPPYMPHSPAG 532  
| | | | |  
RESULT 6  
US-08-818-112-111  
/ Sequence 111, Application US/08818112  
/ Patent No. 6290969  
/ GENERAL INFORMATION:  
/ APPLICANT: Reed, Steven G.  
/ APPLICANT: Skelky, Yasir A.W.  
/ APPLICANT: Dillon, Davin C.  
/ APPLICANT: Campos-Neto, Antonio  
/ APPLICANT: Houghton, Raymond  
/ APPLICANT: Vedvick, Thomas S.  
/ APPLICANT: Twardzik, Daniel R.  
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
/ TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
/ NUMBER OF SEQUENCES: 153  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: SEED and BERRY LLP  
/ STREET: 6300 Columbia Center, 701 Fifth Avenue  
/ CITY: Seattle  
/ STATE: Washington  
/ COUNTRY: USA  
/ ZIP: 98104-7092  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patentin Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/818,112  
/ FILING DATE: 13-MAR-1997  
/ CLASSIFICATION: 424  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Maki, David J.  
/ REGISTRATION NUMBER: 31,392  
/ REFERENCE/DOCKET NUMBER: 210121.411C6  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (206) 622-4900  
/ TELEFAX: (206) 682-6031  
/ INFORMATION FOR SEQ ID NO: 111:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 396 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS:  
/ TOPOLOGY: linear  
US-08-818-112-111  
Query Match 84.8%; Score 1652.5; DB 3; Length 396;

Best Local Similarity 84.9%; Pred. No. 1.3e-129;  
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;  
QY 1 MPDFGLPPEINISARRYAGPGSASLVAAQOMDSVASDLFSAASASQSYVMGLTVGSSWG 60  
| | | | |  
Db 1 VDFGLPPEINISARRYAGPGSASLVAAQOMDSVASDLFSAASASQSYVMGLTVGSSWG 60  
| | | | |  
QY 61 SSAGLWVAASPVVAMSVTGAQELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
| | | | |  
Db 61 SSAGLWVAASPVVAMSVTGAQELTAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
| | | | |  
QY 121 LATNLGONTPAIAVNEAEYEGEMWADAAMFGVAATAATATATLLPFEAEEMTSAGG 180  
| | | | |  
Db 121 LATNLGONTPAIAVNEAEYEGEMWADAAMFGVAATAATATATLLPFEAEEMTSAGG 180  
| | | | |  
QY 181 LLEQAAVVEASDTAAANQMLNNVPALQOLAPTOGTTSSKLGLMKTVSHRSPISN 240  
| | | | |  
Db 181 LLEQAAVVEASDTAAANQMLNNVPALQOLAPTOGTTSSKLGLMKTVSHRSPISN 240  
| | | | |  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVTAQNGVRAMSS---LGSSLS 296  
| | | | |  
Db 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVTAQNGVRAMSS---LGSSLS 296  
| | | | |  
QY 297 GSSGLGAGVAAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGQMLG 356  
| | | | |  
Db 300 GSSGLGAGVAAANIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGQMLG 359  
| | | | |  
QY 357 GIPVGMGARAG--GGLSGVLVPPPYMPHSPAG 391  
| | | | |  
Db 360 GIPVGMGARAG--GGLSGVLVPPPYMPHSPAG 396  
| | | | |  
RESULT 7  
US-08-818-111-106  
/ Sequence 106, Application US/08818111  
/ Patent No. 6338852  
/ GENERAL INFORMATION:  
/ APPLICANT: Reed, Steven G.  
/ APPLICANT: Skelky, Yasir A.W.  
/ APPLICANT: Dillon, Davin C.  
/ APPLICANT: Campos-Neto, Antonio  
/ APPLICANT: Houghton, Raymond  
/ APPLICANT: Vedvick, Thomas S.  
/ APPLICANT: Twardzik, Daniel R.  
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
/ TITLE OF INVENTION: TUBERCULOSIS  
/ NUMBER OF SEQUENCES: 148  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: SEED and BERRY LLP  
/ STREET: 6300 Columbia Center, 701 Fifth Avenue  
/ CITY: Seattle  
/ STATE: Washington  
/ COUNTRY: USA  
/ ZIP: 98104-7092  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patentin Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/818,111  
/ FILING DATE: 13-MAR-1997  
/ CLASSIFICATION: 424  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Maki, David J.  
/ REGISTRATION NUMBER: 31,392  
/ REFERENCE/DOCKET NUMBER: 210121.417C6  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (206) 622-4900  
/ TELEFAX: (206) 682-6031  
/ INFORMATION FOR SEQ ID NO: 106:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 396 amino acids  
/ TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear  
US-08-818-111-106

Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
Best Local Similarity 84.9%; Pred. No. 1.3e-129;  
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPEINRARNYAGPGSASLVAAQMDVSVDLFSASASQSVYWGTLTWSWG 60  
DB 1 VDFGALPEINRARNYAGPGSASLVAAQMDVSVDLFSASASQSVYWGTLTWSWG 60  
QY 61 SSAGLWVAASPVVAMSVTAGOAEITAAQVRAAAAYETAYGLTVPPIAENRAELMT 120  
DB 61 SSAGLWVAASPVVAMSVTAGOAEITAAQVRAAAAYETAYGLTVPPIAENRAELMT 120  
QY 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGYAATATATALLPEEAPEMTSAGG 180  
DB 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGYAATATATALLPEEAPEMTSAGG 180  
QY 181 LLEQAAVEASDTAAANOLMNNVPOALQOAPTOGTTSSKLGKMTVSPHRSPISN 240  
DB 181 LLEQAAVEASDTAAANOLMNNVPOALQOAPTOGTTSSKLGKMTVSPHRSPISN 240  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAONGVRAMSS----LGSSL 296  
DB 241 IVSMNNHVSMTNSGVSMATLHSMKGFAP-AAQAQVETAONGVQAMSSLGSSL 299  
QY 297 GSSGLGGVAAANLGRAASVGSLSVPOMAAAQAVTPPARALPLTSLTSAERPGOMLG 356  
DB 300 GSSGLGAGVAAANLGRAASVGSLSVPOMAAAQAVTPPARALPLTSLTSAERPGOMLG 359  
QY 357 GLPVGQMGARAG--GGLSGLRVPPRPYPVPHSPAG 391  
DB 360 GLPLGQLTNSGGGFGVSNALRMPRAVYMPRVPAAG 396

## RESULT 8

US-09-056-556-111  
Sequence 111, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-Apr-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 111:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids

TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-056-556-111

Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
Best Local Similarity 84.9%; Pred. No. 1.3e-129;  
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPEINRARNYAGPGSASLVAAQMDVSVDLFSASASQSVYWGTLTWSWG 60  
DB 1 VDFGALPEINRARNYAGPGSASLVAAQMDVSVDLFSASASQSVYWGTLTWSWG 60  
QY 61 SSAGLWVAASPVVAMSVTAGOAEITAAQVRAAAAYETAYGLTVPPIAENRAELMT 120  
DB 61 SSAGLWVAASPVVAMSVTAGOAEITAAQVRAAAAYETAYGLTVPPIAENRAELMT 120  
QY 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGYAATATATALLPEEAPEMTSAGG 180  
DB 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGYAATATATALLPEEAPEMTSAGG 180  
QY 181 LLEQAAVEASDTAAANOLMNNVPOALQOAPTOGTTSSKLGKMTVSPHRSPISN 240  
DB 181 LLEQAAVEASDTAAANOLMNNVPOALQOAPTOGTTSSKLGKMTVSPHRSPISN 240  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAONGVRAMSS----LGSSL 296  
DB 241 IVSMNNHVSMTNSGVSMATLHSMKGFAP-AAQAQVETAONGVQAMSSLGSSL 299  
QY 297 GSSGLGGVAAANLGRAASVGSLSVPOMAAAQAVTPPARALPLTSLTSAERPGOMLG 356  
DB 300 GSSGLGAGVAAANLGRAASVGSLSVPOMAAAQAVTPPARALPLTSLTSAERPGOMLG 359  
QY 357 GLPVGQMGARAG--GGLSGLRVPPRPYPVPHSPAG 391  
DB 360 GLPLGQLTNSGGGFGVSNALRMPRAVYMPRVPAAG 396

## RESULT 9

US-09-072-596-106  
Sequence 106, Application US/09072596  
Patent No. 6458366  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 106:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 396 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-09-072-596-106

Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
 Best Local Similarity 84.9%; Pred. No. 1.3e-129;  
 Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MNDGALPEPINSARMYAGGASLVAAQMDVSADLFSASAFQSVWGLTVGSWIG 60  
 DB 1 VDPGALPEPINSARMYAGGASLVAAQMDVSADLFSASAFQSVWGLTVGSWIG 60  
 QY 61 SSAGLWVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
 DB 61 SSAGLWVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
 QY 121 LIATNLGONTPAIVAEYEGEMWADAAAMFGYAAATATATATLLPPEAPEMTSAGG 180  
 DB 121 LIATNLGONTPAIVAEYEGEMWADAAAMFGYAAATATATATLLPPEAPEMTSAGG 180  
 QY 181 LLEQAAVEASPTAANQNMNVPOALQOAPTOGTPTSSKLGIMKTVSPHRSPTSN 240  
 DB 181 LLEQAAVEASPTAANQNMNVPOALQOAPTOGTPTSSKLGIMKTVSPHRSPTSN 240  
 QY 241 MVSANNNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTAONGVRAMSS---LGSSL 296  
 DB 241 IVSMNNHVSMTSGVSMNTTLLSMLKGFAPAAAQAQVOTAONGVRAMSSLGSSL 296  
 QY 297 GSSGLGGVAVANIGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERPGQMTG 356  
 DB 300 GSSGLGGVAVANIGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERPGHMLG 359  
 QY 357 GLPVGQMGARAG--GGLSVLRVPRPYWPHSPAAG 391  
 DB 360 GLPVGQLTNSGGFGVSNALRMPPRAYVMPRVPAAG 396

# RESULT 10

US-08-818-112-109  
 Sequence 109, Application US/08818112  
 Patent No. 6290969

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Campos-Neco, Antonio  
 APPLICANT: Houghton, Raymond  
 APPLICANT: Vedvick, Thomas S.  
 APPLICANT: Twardzik, Daniel R.  
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
 TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
 NUMBER OF SEQUENCES: 153  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED AND BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/818,112  
 FILING DATE: 13-MAR-1997  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mark, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.411C6  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 109:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 359 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-818-112-109

Query Match 76.3%; Score 1486.5; DB 3; Length 359;  
 Best Local Similarity 84.2%; Pred. No. 7.2e-116;  
 Matches 303; Conservative 16; Mismatches 36; Indels 5; Gaps 2;

QY 1 MNDGALPEPINSARMYAGGASLVAAQMDVSADLFSASAFQSVWGLTVGSWIG 60  
 DB 1 VDPGALPEPINSARMYAGGASLVAAQMDVSADLFSASAFQSVWGLTVGSWIG 60  
 QY 61 SSAGLWVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
 DB 61 SSAGLWVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
 QY 121 LIATNLGONTPAIVAEYEGEMWADAAAMFGYAAATATATATLLPPEAPEMTSAGG 180  
 DB 121 LIATNLGONTPAIVAEYEGEMWADAAAMFGYAAATATATATLLPPEAPEMTSAGG 180  
 QY 181 LLEQAAVEASPTAANQNMNVPOALQOAPTOGTPTSSKLGIMKTVSPHRSPTSN 240  
 DB 181 LLEQAAVEASPTAANQNMNVPOALQOAPTOGTPTSSKLGIMKTVSPHRSPTSN 240  
 QY 241 MVSANNNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTAONGVRAMSS---LGSSL 296  
 DB 241 VSSIANNNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTAONGVRAMSSLGSSL 299  
 QY 297 GSSGLGGVAVANIGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERPGQMTG 356  
 DB 300 GSSGLGGVAVANIGRAASVGSLSVPPAMAAANQAVTPAARALPLTSLTSAERPGHMLG 359

# RESULT 11

US-08-818-111-104  
 Sequence 104, Application US/08818111  
 Patent No. 6338852

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Campos-Neco, Antonio  
 APPLICANT: Houghton, Raymond  
 APPLICANT: Vedvick, Thomas S.  
 APPLICANT: Twardzik, Daniel R.  
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
 TITLE OF INVENTION: AND METHODS FOR IMMUNOTHERAPY  
 NUMBER OF SEQUENCES: 148  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED AND BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/06/818,111  
;/ FILING DATE: 13-MAR-1997  
;/ CLASSIFICATION: 424  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Maki, David J.  
;/ REGISTRATION NUMBER: 31,392  
;/ REFERENCE/DOCKET NUMBER: 210121.417C6  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (206) 622-4900  
;/ TELEFAX: (206) 682-6031  
;/ INFORMATION FOR SEQ ID NO: 104:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 359 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS:  
;/ TOPOLOGY: linear  
;/ US-08-818-111-104

Query Match 76.3%; Score 1486.5; DB 4; Length 359;  
Best Local Similarity 84.2%; Pred. No. 7.2e-116;  
Matches 303; Conservative 16; Mismatches 36; Indels 5; Gaps 2;

QY 1 MVDGALPEINSARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 VVDFGALPEINSARMYAGPGSASLVAAAKMDSVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLMAAASPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPIVIAENRAELMT 120  
DB 61 SSAGLMAAASPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPIVIAENRAELMT 120  
QY 121 LIAATNLGONTPAIAVNEAEYEGEMWAODAAAFGYAAATATATATLLPPEEAPEMTSAG 180  
DB 121 LIAATNLGONTPAIAVNEAEYEGEMWAODAAAFGYAAATATATATLLPPEEAPEMTSAG 180  
QY 181 LLEQAAVEASDTAAANQLMNNVPQALQOLAOPGTGTPSSKLGKMTVSPHRSPIEN 240  
DB 181 LLEQAAVEASDTAAANQLMNNVPQALQOLAOPGTGTPSSKLGKMTVSPHRSPIEN 240  
QY 241 WVSMMNNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 296  
DB 241 WVSMMNNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 296  
QY 241 VSSIANNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 299  
DB 241 VSSIANNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 299  
QY 297 GSSGLGGGVAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPOMIG 356  
DB 300 GSSGLGGGVAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPOMIG 359

RESULT 12  
US-09-056-556-109  
; Sequence 109, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998

;/ CLASSIFICATION:  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Maki, David J.  
;/ REGISTRATION NUMBER: 31,392  
;/ REFERENCE/DOCKET NUMBER: 210121.457  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (206) 622-4900  
;/ TELEFAX: (206) 682-6031  
;/ INFORMATION FOR SEQ ID NO: 109:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 359 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS:  
;/ TOPOLOGY: linear  
;/ US-09-056-556-109

Query Match 76.3%; Score 1486.5; DB 4; Length 359;  
Best Local Similarity 84.2%; Pred. No. 7.2e-116;  
Matches 303; Conservative 16; Mismatches 36; Indels 5; Gaps 2;

QY 1 MVDGALPEINSARMYAGPGSASLVAAQMDVSASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 VVDFGALPEINSARMYAGPGSASLVAAAKMDSVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLMAAASPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPIVIAENRAELMT 120  
DB 61 SSAGLMAAASPYVAMSVTAQAELTAAQVRAAAAYETAGLTVPPIVIAENRAELMT 120  
QY 121 LIAATNLGONTPAIAVNEAEYEGEMWAODAAAFGYAAATATATATLLPPEEAPEMTSAG 180  
DB 121 LIAATNLGONTPAIAVNEAEYEGEMWAODAAAFGYAAATATATATLLPPEEAPEMTSAG 180  
QY 181 LLEQAAVEASDTAAANQLMNNVPQALQOLAOPGTGTPSSKLGKMTVSPHRSPIEN 240  
DB 181 LLEQAAVEASDTAAANQLMNNVPQALQOLAOPGTGTPSSKLGKMTVSPHRSPIEN 240  
QY 241 WVSMMNNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 296  
DB 241 WVSMMNNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 296  
QY 241 VSSIANNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 299  
DB 241 VSSIANNHSMNTSGVSMNTLTSSMLKGFAPAAQAAVOTAQNGVRAMSS---LGSSL 299  
QY 297 GSSGLGGGVAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPOMIG 356  
DB 300 GSSGLGGGVAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPOMIG 359

RESULT 13  
US-09-072-596-104  
; Sequence 104, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twartzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-072-596-104

Query Match 76.3%; Score 1486.5; DB 4; Length 359;  
Best Local Similarity 84.2%; Pred. No. 7.2e-116;  
Matches 303; Conservative 16; Mismatches 36; Indels 5; Gaps 2;

QY 1 MNDPGLPEPINSARMTAGCGSASLVAAQMDVSASDLFSASAFOSVYWGTLVSGWIG 60  
DB 1 VVDFGALPPEINSARMTAGCGSASLVAAQMDVSASDLFSASAFOSVYWGTLVSGWIG 60  
QY 61 SSAGLWAAASPYVAMSVTAAGQELTAQVRAAAVETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWAAASPYVAMSVTAAGQELTAQVRAAAVETAYGLTVPPVIAENRAELMI 120  
QY 121 LIATNLGONTPAVAVEAEYGEVMAADAAAFGYAATATATATLLPFEBAPEMTSAG 180  
DB 121 LIATNLGONTPAVAVEAEYGEVMAADAAAFGYAATATATATLLPFEBAPEMTSAG 180  
QY 181 LLEBAAVEASDTAAANQLMNNVPAQLQOLAPTOGTTSSKLGIMKTVSPHRSPLSN 240  
DB 181 LLEBAAVEASDTAAANQLMNNVPAQLQOLAPTOGTTSSKLGIMKTVSPHRSPLSN 240  
QY 241 MVSANNNHSMSTNGVSMNTTLLSSMLGFPAPAAAQAQVOTAONGVRAMSS---LGSSL 296  
DB 241 VSSIANNNHSMSTNGVSMNTTLLSSMLGFPAPAAAQAQVOTAONGVRAMSS---LGSSL 296  
QY 297 GSSGLGGVAANTGRAASVGSLSVPOAMAAANQVTPPARALPLTSLTSAERGPQGMIG 356  
DB 300 GSSGLGGVAANTGRAASVGSLSVPOAMAAANQVTPPARALPLTSLTSAERGPQGMIG 359

RESULT 14  
US-08-818-112-91  
Sequence 91, Application US/08818112  
Patent No. 6290969  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 153  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,112  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-112-91

Query Match 60.9%; Score 1187; DB 3; Length 263;  
Best Local Similarity 99.6%; Pred. No. 3.9e-91;  
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 74 VAMMSYTAGQELTAAQVRAAAVETAYGLTVPPVIAENRAELMI,LIATNLGONTPA 133  
DB 1 VAMMSYTAGQELTAAQVRAAAVETAYGLTVPPVIAENRAELMI,LIATNLGONTPA 60  
QY 134 IAVNEAEYGEVMAADAAAFGYAATATATATLLPFEBAPEMTSAGLLEQAAAVEASD 193  
DB 61 IAVNEAEYGEVMAADAAAFGYAATATATATLLPFEBAPEMTSAGLLEQAAAVEASD 120  
QY 194 TAAANQLMNNVPAQLQOLAPTOGTTSSKLGIMKTVSPHRSPLSNMVSANNNHSMSTN 253  
DB 121 TAAANQLMNNVPAQLQOLAPTOGTTSSKLGIMKTVSPHRSPLSNMVSANNNHSMSTN 180  
QY 254 SGVSMNTTLLSSMLGFPAPAAAQAQVOTAONGVRAMSSLGSSGLGGVAANTGRAA 313  
DB 181 SGVSMNTTLLSSMLGFPAPAAAQAQVOTAONGVRAMSSLGSSGLGGVAANTGRAA 240  
QY 314 SV 315  
DB 241 SV 242

RESULT 15  
US-08-818-111-92  
Sequence 92, Application US/08818111  
Patent No. 6338852  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/08/818,111  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 92:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-111-92

Query Match 60.9%; Score 1187; DB 4; Length 263;  
Best Local Similarity 99.6%; Pred. No. 3.9e-91;  
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	74	VAMSVTAQGAELTAAQVRAAAAYETAVGLTVPPVIAENRAELMILIAITNLLGNTPA	133
DB	1	VAMSVTAQGAELTAAQVRAAAAYETAVGLTVPPVIAENRAELMILIAITNLLGNTPA	60
QY	134	IAVNEAEVGEEMWAQDAAMFGYAAATATATATLLPPEEAPEMTSAGGLEQAAAVEASD	193
DB	61	IAVNEAEVGEEMWAQDAAMFGYAAATATATATLLPPEEAPEMTSAGGLEQAAAVEASD	120
QY	194	TAAANQLMNNVPOALQOALQPTGTTSSKLGGLMKTVPHSRPIINMTSMANNHSMTN	253
DB	121	TAAANQLMNNVPOALQOALQPTGTTSSKLGGLMKTVPHSRPIINMTSMANNHSMTN	180
QY	254	SGVSMNTNLTSSMLKGFAPAAAAQAVOTAAQNGVRANSLGSSLGSGGVAANLGRAA	313
DB	181	SGVSMNTNLTSSMLKGFAPAAAAQAVOTAAQNGVRANSLGSSLGSGGVAANLGRAA	240
QY	314	SV 315	
DB	241	SV 242	

Search completed: November 21, 2003, 17:06:38  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:04:15 ; Search time 173 Seconds  
(without alignments)  
2056.522 Million cell updates/sec

Title: US-09-724-685-107  
Perfect score: 1949

Sequence: 1 MWDFGALPPEINSARMYACP.....SGVLRVPPRPYVMPHSPAAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : pending\_Patents\_AA\_Main:\*

1:	/cgn2_6/prodata/1/paa/PCTTS_COMB.pcp.*
2:	/cgn2_6/prodata/1/paa/US06_COMB.pcp.*
3:	/cgn2_6/prodata/1/paa/US07_COMB.pcp.*
4:	/cgn2_6/prodata/1/paa/US08_COMB.pcp.*
5:	/cgn2_6/prodata/1/paa/US09_COMB.pcp.*
6:	/cgn2_6/prodata/1/paa/US092_COMB.pcp.*
7:	/cgn2_6/prodata/1/paa/US083_COMB.pcp.*
8:	/cgn2_6/prodata/1/paa/US084_COMB.pcp.*
9:	/cgn2_6/prodata/1/paa/US085_COMB.pcp.*
10:	/cgn2_6/prodata/1/paa/US086_COMB.pcp.*
11:	/cgn2_6/prodata/1/paa/US087_COMB.pcp.*
12:	/cgn2_6/prodata/1/paa/US088_COMB.pcp.*
13:	/cgn2_6/prodata/1/paa/US098_COMB.pcp.*
14:	/cgn2_6/prodata/1/paa/US090_COMB.pcp.*
15:	/cgn2_6/prodata/1/paa/US091_COMB.pcp.*
16:	/cgn2_6/prodata/1/paa/US092_COMB.pcp.*
17:	/cgn2_6/prodata/1/paa/US093_COMB.pcp.*
18:	/cgn2_6/prodata/1/paa/US094_COMB.pcp.*
19:	/cgn2_6/prodata/1/paa/US095_COMB.pcp.*
20:	/cgn2_6/prodata/1/paa/US096_COMB.pcp.*
21:	/cgn2_6/prodata/1/paa/US097A_COMB.pcp.*
22:	/cgn2_6/prodata/1/paa/US097B_COMB.pcp.*
23:	/cgn2_6/prodata/1/paa/US098_COMB.pcp.*
24:	/cgn2_6/prodata/1/paa/US099A_COMB.pcp.*
25:	/cgn2_6/prodata/1/paa/US099B_COMB.pcp.*
26:	/cgn2_6/prodata/1/paa/US100_COMB.pcp.*
27:	/cgn2_6/prodata/1/paa/US101_COMB.pcp.*
28:	/cgn2_6/prodata/1/paa/US102_COMB.pcp.*
29:	/cgn2_6/prodata/1/paa/US103_COMB.pcp.*
30:	/cgn2_6/prodata/1/paa/US104_COMB.pcp.*
31:	/cgn2_6/prodata/1/paa/US106_COMB.pcp.*
32:	/cgn2_6/prodata/1/paa/US10 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query				
No.	Score	Match	Length	ID	Description
1	1949	100.0	391	1	PCT-US99-03265-102
					Sequence 102, App

2	1949	100.0	331	1	PCT-US99-03268-107	Sequence 107, App
3	1949	100.0	331	10	US-08-656-800-102	Sequence 102, App
4	1949	100.0	331	10	US-08-656-683-107	Sequence 107, App
5	1949	100.0	381	10	US-08-680-573-102	Sequence 102, App
6	1949	100.0	331	10	US-08-680-574-107	Sequence 107, App
7	1949	100.0	331	11	US-08-728-622-102	Sequence 102, App
8	1949	100.0	331	11	US-08-730-510-107	Sequence 107, App
9	1949	100.0	331	13	US-08-942-341-102	Sequence 102, App
10	1949	100.0	331	13	US-08-942-578-107	Sequence 107, App
11	1949	100.0	331	14	US-09-024-753-102	Sequence 102, App
12	1949	100.0	331	14	US-09-024-197-107	Sequence 107, App
13	1949	100.0	331	14	US-09-072-967-107	Sequence 107, App
14	1949	100.0	331	19	US-09-597-786C-8	Sequence 8, App1
15	1949	100.0	331	20	US-09-688-672A-26	Sequence 26, App1
16	1949	100.0	331	21	US-09-724-665-107	Sequence 107, App
17	1949	100.0	331	23	US-09-855-604-107	Sequence 107, App
18	1949	100.0	331	23	US-09-855-604A-107	Sequence 107, App
19	1949	100.0	331	26	US-10-084-883-107	Sequence 107, App
20	1949	100.0	331	26	US-10-098-732A-14	Sequence 14, App
21	1949	100.0	331	27	US-10-199-002-102	Sequence 102, App
22	1949	100.0	334	23	US-09-855-604-109	Sequence 109, App
23	1949	100.0	334	23	US-09-855-604A-109	Sequence 109, App
24	1949	100.0	536	16	US-09-287-8849-26	Sequence 26, App
25	1949	100.0	536	19	US-09-597-796C-10	Sequence 10, App
26	1949	100.0	536	26	US-10-098-732A-20	Sequence 20, App
27	1949	100.0	536	29	US-10-359-460-26	Sequence 26, App
28	1949	100.0	600	16	US-09-287-8849-22	Sequence 22, App
29	1949	100.0	600	29	US-10-355-460-22	Sequence 22, App
30	1949	100.0	723	1	PCT-US03-04903-2	Sequence 2, App1
31	1949	100.0	729	1	PCT-US00-27652-12	Sequence 12, App1
32	1949	100.0	729	1	PCT-US03-04903-21	Sequence 21, App1
33	1949	100.0	729	1	PCT-US03-04903-22	Sequence 22, App1
34	1949	100.0	729	20	US-09-684-215A-12	Sequence 12, App
35	1949	100.0	729	26	US-10-098-732A-18	Sequence 18, App
36	1949	100.0	780	20	US-09-688-672A-58	Sequence 58, App
37	1949	100.0	780	20	US-09-688-672A-60	Sequence 60, App
38	1949	100.0	811	20	US-09-688-672A-62	Sequence 62, App
39	1949	100.0	813	1	PCT-US03-04903-15	Sequence 15, App1
40	1949	100.0	825	1	PCT-US03-04903-14	Sequence 14, App1
41	1949	100.0	875	1	PCT-US03-04903-13	Sequence 13, App1
42	1949	100.0	930	1	PCT-US03-04903-12	Sequence 12, App1
43	1949	100.0	930	26	US-10-098-732A-65	Sequence 65, App
44	1949	100.0	1010	1	PCT-US03-04903-4	Sequence 4, App1
45	1949	100.0	1016	1	PCT-US03-04903-18	Sequence 18, App1

## ALIGNMENTS

RESULT 1  
PCT-US99-03265-102  
Sequence 102. Application PC/RUS9903265  
GENERAL INFORMATION:  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
TUBERCULOSIS  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds, LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US99/03265  
FILING DATE: 17-FEB-1999  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/024,753  
FILING DATE: 18-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Laura A. Coruzzi  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 9532-0023-228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US99-03265-102

Query Match 100.0%; Score 1949; DB 1; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARYAGPGSASLVAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPEINSAARYAGPGSASLVAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAAGLTVPPIAENRAELMT 120  
DB 61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAAGLTVPPIAENRAELMT 120  
QY 121 LIATNLGONTPAIAVNEAEYGEEMQAODAAAFGVAATAATATATLLPFEAEEMTSAGG 180  
DB 121 LIATNLGONTPAIAVNEAEYGEEMQAODAAAFGVAATAATATATLLPFEAEEMTSAGG 180  
QY 181 LLEQAAAVEASDPTAANQIMNNVPQALQLOAOPGTTTSSKLGIMKTVSFHRSPISN 240  
DB 181 LLEQAAAVEASDPTAANQIMNNVPQALQLOAOPGTTTSSKLGIMKTVSFHRSPISN 240  
QY 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVQTAONGVAMSSLGSSSG 300  
DB 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVQTAONGVAMSSLGSSSG 300  
QY 301 LGGVAANIGRAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERPGQMLGSLPV 360  
DB 301 LGGVAANIGRAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERPGQMLGSLPV 360  
QY 361 GQMGARAGGLSGVLRVPPRYVMPHSPAAG 391  
DB 361 GQMGARAGGLSGVLRVPPRYVMPHSPAAG 391

RESULT 2  
PCT-US99-03268-107  
Sequence 107, Application PC/TUS9903268  
GENERAL INFORMATION:  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds, LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US99/03268  
FILING DATE: 17-FEB-1999  
CLASSIFICATION: 406  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/025,197  
FILING DATE: 18-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 9532-0013-228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US99-03268-107

Query Match 100.0%; Score 1949; DB 1; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARYAGPGSASLVAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPEINSAARYAGPGSASLVAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAAGLTVPPIAENRAELMT 120  
DB 61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAAGLTVPPIAENRAELMT 120  
QY 121 LIATNLGONTPAIAVNEAEYGEEMQAODAAAFGVAATAATATATLLPFEAEEMTSAGG 180  
DB 121 LIATNLGONTPAIAVNEAEYGEEMQAODAAAFGVAATAATATATLLPFEAEEMTSAGG 180  
QY 181 LLEQAAAVEASDPTAANQIMNNVPQALQLOAOPGTTTSSKLGIMKTVSFHRSPISN 240  
DB 181 LLEQAAAVEASDPTAANQIMNNVPQALQLOAOPGTTTSSKLGIMKTVSFHRSPISN 240  
QY 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVQTAONGVAMSSLGSSSG 300  
DB 241 MWSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVQTAONGVAMSSLGSSSG 300  
QY 301 LGGVAANIGRAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERPGQMLGSLPV 360  
DB 301 LGGVAANIGRAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERPGQMLGSLPV 360  
QY 361 GQMGARAGGLSGVLRVPPRYVMPHSPAAG 391  
DB 361 GQMGARAGGLSGVLRVPPRYVMPHSPAAG 391

RESULT 3  
US-08-658-800-102  
Sequence 102, Application US/08658800  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 123  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/658,800  
FILING DATE: 05-JUN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Naki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-658-800-102

Query Match 100.0%; Score 1949; DB 10; Length 391;

Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEEINSARVAGPGSASLVAAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPEEINSARVAGPGSASLVAAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLMVAAAPYVAMVSTAGQAEITAAQVRAAAYETAAGLTVPPIVIAENRAELMI 120  
DB 61 SSAGLMVAAAPYVAMVSTAGQAEITAAQVRAAAYETAAGLTVPPIVIAENRAELMI 120  
QY 121 LIATNLGONTPAIVNEAEYEGEMAAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
DB 121 LIATNLGONTPAIVNEAEYEGEMAAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
QY 181 LLEQAAVEASDTAAANOQMNVPQALQLOPTGTTSSKLGIMKTVPBHRPISN 240  
DB 181 LLEQAAVEASDTAAANOQMNVPQALQLOPTGTTSSKLGIMKTVPBHRPISN 240  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSSG 300  
DB 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSSG 300  
QY 301 LGGVAAANLGRAASVGSISVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
DB 301 LGGVAAANLGRAASVGSISVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAA 391  
DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAA 391

RESULT 4

US-08-659-683-107

Sequence 107, Application US/08659683  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/659,683  
FILING DATE: 05-JUN-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Naki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-659-683-107

Query Match 100.0%; Score 1949; DB 10; Length 391;

Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEEINSARVAGPGSASLVAAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPEEINSARVAGPGSASLVAAAQMDVVASDLFSAASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLMVAAAPYVAMVSTAGQAEITAAQVRAAAYETAAGLTVPPIVIAENRAELMI 120  
DB 61 SSAGLMVAAAPYVAMVSTAGQAEITAAQVRAAAYETAAGLTVPPIVIAENRAELMI 120  
QY 121 LIATNLGONTPAIVNEAEYEGEMAAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
DB 121 LIATNLGONTPAIVNEAEYEGEMAAQDAAMFGYAAATATATATLLPPEEAPEMTSAG 180  
QY 181 LLEQAAVEASDTAAANOQMNVPQALQLOPTGTTSSKLGIMKTVPBHRPISN 240  
DB 181 LLEQAAVEASDTAAANOQMNVPQALQLOPTGTTSSKLGIMKTVPBHRPISN 240  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSSG 300  
DB 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSSG 300  
QY 301 LGGVAAANLGRAASVGSISVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
DB 301 LGGVAAANLGRAASVGSISVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAA 391  
DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAA 391

RESULT 5

US-08-680-573-102

Sequence 102, Application US/08680573  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/680,573  
FILING DATE: 12-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

Query Match 100.0%; Score 1949; DB 10; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARVAGPGSASLVAAQMDVSADLFSASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPEINSARVAGPGSASLVAAQMDVSADLFSASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
QY 121 LIATNLGONTPIAIVNEAEYGEEMWQDAAMFGVAAATATATATLLPFEAEEMTSAGG 180  
DB 121 LIATNLGONTPIAIVNEAEYGEEMWQDAAMFGVAAATATATATLLPFEAEEMTSAGG 180  
QY 181 LLEQAAVEASDPTAAANQIMNNVPOALQLOPTGTTSSKLGIMKTVPHRSPISN 240  
DB 181 LLEQAAVEASDPTAAANQIMNNVPOALQLOPTGTTSSKLGIMKTVPHRSPISN 240  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 300  
DB 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 300  
QY 301 LGGGVAAANGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
DB 301 LGGGVAAANGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLVPPRPVYMPHSPAG 391  
DB 361 GQMGARAGGGLSGVLVPPRPVYMPHSPAG 391

RESULT 6  
US-08-680-574-107  
Sequence 107, Application US/08680574  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
APPLICANT: Campos-Neto, Antonio  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 133  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington

COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/680,574  
FILING DATE: 12-JUL-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

Query Match 100.0%; Score 1949; DB 10; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARVAGPGSASLVAAQMDVSADLFSASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPEINSARVAGPGSASLVAAQMDVSADLFSASAFQSVVWGLTVGSWIG 60  
QY 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWVAASPVAAMSVTAQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
QY 121 LIATNLGONTPIAIVNEAEYGEEMWQDAAMFGVAAATATATATLLPFEAEEMTSAGG 180  
DB 121 LIATNLGONTPIAIVNEAEYGEEMWQDAAMFGVAAATATATATLLPFEAEEMTSAGG 180  
QY 181 LLEQAAVEASDPTAAANQIMNNVPOALQLOPTGTTSSKLGIMKTVPHRSPISN 240  
DB 181 LLEQAAVEASDPTAAANQIMNNVPOALQLOPTGTTSSKLGIMKTVPHRSPISN 240  
QY 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 300  
DB 241 MYSMANNHSMNTSGVSMNTTSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSG 300  
QY 301 LGGGVAAANGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
DB 301 LGGGVAAANGRAASVGSLSVPOAMAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
QY 361 GQMGARAGGGLSGVLVPPRPVYMPHSPAG 391  
DB 361 GQMGARAGGGLSGVLVPPRPVYMPHSPAG 391

RESULT 7  
US-08-729-622-102  
Sequence 102, Application US/08729622  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas H.  
APPLICANT: Twardzik, David R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
TITLE OF INVENTION: TUBERCULOSIS  
NUMBER OF SEQUENCES: 132

```

CORRESPONDENCE ADDRESS:
ADDRESSER: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,622
FILING DATE: 11-OCT-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C5
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 391 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-729-622-102

```

Query Match 100.0%; Score 1949; DB 11; Length 391;

Best Local Similarity 100.0%; Pred. No. 1.7e-155; Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVWGTLTVGSWIG 60
DB 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVWGTLTVGSWIG 60
QY 61 SSAGLMTVAASPYVAMSTYTAQAELTAAQVRAAAAYETAYGLTVPPIVIAENRAELMI 120
DB 61 SSAGLMTVAASPYVAMSTYTAQAELTAAQVRAAAAYETAYGLTVPPIVIAENRAELMI 120
QY 121 LATNLGONTPIAVNEAEYGEEMADAAAFGVAAATATATATLLPREBAPEMTSAG 180
DB 121 LATNLGONTPIAVNEAEYGEEMADAAAFGVAAATATATATLLPREBAPEMTSAG 180
QY 181 LLEQAAAVEASDTAAANOLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 181 LLEQAAAVEASDTAAANOLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
QY 241 MYSMANNHSMNTSGVSMNTLSSMLKGPAPAAAQAQVTAQNGVRAMSLGSSIGSSG 300
DB 241 MYSMANNHSMNTSGVSMNTLSSMLKGPAPAAAQAQVTAQNGVRAMSLGSSIGSSG 300
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPOMLGGLPV 360
DB 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPOMLGGLPV 360
QY 361 GQMGARAGGSLGVLRVPPRPVPMHSPPAAG 391
DB 361 GQMGARAGGSLGVLRVPPRPVPMHSPPAAG 391

```

RESULT 8  
US-08-730-510-107

; Sequence 107, Application US/08730510

; GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond

```

APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 137
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/730,510
FILING DATE: 27-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C5
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 391 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-730-510-107

```

Query Match 100.0%; Score 1949; DB 11; Length 391;

Best Local Similarity 100.0%; Pred. No. 1.7e-155; Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVWGTLTVGSWIG 60
DB 1 MVDGALPEINSARMYAGPGSASLVAAAQMDVSADLFSASAFQSVVWGTLTVGSWIG 60
QY 61 SSAGLMTVAASPYVAMSTYTAQAELTAAQVRAAAAYETAYGLTVPPIVIAENRAELMI 120
DB 61 SSAGLMTVAASPYVAMSTYTAQAELTAAQVRAAAAYETAYGLTVPPIVIAENRAELMI 120
QY 121 LATNLGONTPIAVNEAEYGEEMADAAAFGVAAATATATATLLPREBAPEMTSAG 180
DB 121 LATNLGONTPIAVNEAEYGEEMADAAAFGVAAATATATATLLPREBAPEMTSAG 180
QY 181 LLEQAAAVEASDTAAANOLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
DB 181 LLEQAAAVEASDTAAANOLMNNVPOALQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240
QY 241 MYSMANNHSMNTSGVSMNTLSSMLKGPAPAAAQAQVTAQNGVRAMSLGSSIGSSG 300
DB 241 MYSMANNHSMNTSGVSMNTLSSMLKGPAPAAAQAQVTAQNGVRAMSLGSSIGSSG 300
QY 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPOMLGGLPV 360
DB 301 LGGGVAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERGPOMLGGLPV 360
QY 361 GQMGARAGGSLGVLRVPPRPVPMHSPPAAG 391
DB 361 GQMGARAGGSLGVLRVPPRPVPMHSPPAAG 391

```

RESULT 9  
US-08-942-341-102

; Sequence 102, Application US/08942341

; GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vardzik, Thomas S.  
APPLICANT: Lodes, Michael J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/942,341  
FILING DATE: 01-OCT-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C7  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-942-341-102

Query Match 100.0%; Score 1949; DB 13; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVDGALPEINSAARYAGPGSASLVAAQOMDSVASDLSAASAFQSVVWGLTVGSWIG 60  
1 MVDGALPEINSAARYAGPGSASLVAAQOMDSVASDLSAASAFQSVVWGLTVGSWIG 60  
61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMT 120  
61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMT 120  
121 LIATNLGONTPLAIVNEAEYEGEMWQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
121 LIATNLGONTPLAIVNEAEYEGEMWQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
121 LIATNLGONTPLAIVNEAEYEGEMWQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
181 LLEQAAVEASPTAANQLMNNVPOLQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240  
181 LLEQAAVEASPTAANQLMNNVPOLQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240  
181 LLEQAAVEASPTAANQLMNNVPOLQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240  
241 MWSMANNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTPAONGVRAVMSLSSLSGSSG 300  
241 MWSMANNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTPAONGVRAVMSLSSLSGSSG 300  
301 LGGGVAANTIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGLPV 360  
301 LGGGVAANTIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGLPV 360  
361 GQMGARAGGLSGVLRVPRPYVMPHSPAG 391  
361 GQMGARAGGLSGVLRVPRPYVMPHSPAG 391

RESULT 10  
US-08-942-578-107  
Sequence 107, Application US/08942578  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vardzik, Thomas S.  
APPLICANT: Lodes, Michael J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 214  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/942,578  
FILING DATE: 01-OCT-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C7  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-942-578-107

Query Match 100.0%; Score 1949; DB 13; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVDGALPEINSAARYAGPGSASLVAAQOMDSVASDLSAASAFQSVVWGLTVGSWIG 60  
1 MVDGALPEINSAARYAGPGSASLVAAQOMDSVASDLSAASAFQSVVWGLTVGSWIG 60  
61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMT 120  
61 SSAGLWVAASPVAWMSVTAGQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMT 120  
121 LIATNLGONTPLAIVNEAEYEGEMWQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
121 LIATNLGONTPLAIVNEAEYEGEMWQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
121 LIATNLGONTPLAIVNEAEYEGEMWQDAAMFGYAAATATATATLLPFEAEEMTSAGG 180  
181 LLEQAAVEASPTAANQLMNNVPOLQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240  
181 LLEQAAVEASPTAANQLMNNVPOLQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240  
181 LLEQAAVEASPTAANQLMNNVPOLQOLAOPTGTTSSKLGIMKTVPSPHRSPISN 240  
241 MWSMANNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTPAONGVRAVMSLSSLSGSSG 300  
241 MWSMANNHSMNTSGVSMNTTLLSMLKGFAPAAAQAQVOTPAONGVRAVMSLSSLSGSSG 300  
301 LGGGVAANTIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGLPV 360  
301 LGGGVAANTIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGLPV 360



Db 301 LGGVAAANGRAASVGSLSVPQMAAANOAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
Qy 361 GOMGARAGGSLGVLRVPPRPVMPHSPAAG 391  
Db 361 GOMGARAGGSLGVLRVPPRPVMPHSPAAG 391

RESULT 11  
US-09-024-753-102  
; Sequence 102; Application US/09024753  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; TUBERCULOSIS  
; NUMBER OF SEQUENCES: 236  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/024,753  
; FILING DATE: 18-FEB-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 102:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-024-753-102

Query Match 100.0%; Score 1949; DB 14; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDFGALPPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFQSVYWGILTVGSSWG 60  
Db 1 MYDFGALPPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFQSVYWGILTVGSSWG 60  
Qy 61 SSAGLWVAASPYVAMSVTAAGAEITTAQOVVAAAAYETAAGLTVPPVIAENRAELMI 120  
Db 61 SSAGLWVAASPYVAMSVTAAGAEITTAQOVVAAAAYETAAGLTVPPVIAENRAELMI 120  
Qy 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGVAAATATATATLLPPEEAPEMTSAG 180  
Db 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGVAAATATATATLLPPEEAPEMTSAG 180  
Qy 181 LLEQAAVAEASSTAANQAMNVPAOLQOLAOPITGTTSSSLGGLMKTIVSPHRSPISN 240  
Db 181 LLEQAAVAEASSTAANQAMNVPAOLQOLAOPITGTTSSSLGGLMKTIVSPHRSPISN 240  
Qy 241 MYSMANNHSMNTSVGSMNTTILSSMLKGFAPAAAQAQVTAQNGVRAANSISGSSG 300

Db 241 MYSMANNHSMNTSVGSMNTTILSSMLKGFAPAAAQAQVTAQNGVRAANSISGSSG 300  
Qy 301 LGGVAAANGRAASVGSLSVPQMAAANOAVTPAARALPLTSLTSAERPGQMLGGLPV 360  
Db 301 LGGVAAANGRAASVGSLSVPQMAAANOAVTPAARALPLTSLTSAERPGQMLGGLPV 360

RESULT 12  
US-09-025-197-107  
; Sequence 107; Application US/09025197  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TUBERCULOSIS  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/025,197  
; FILING DATE: 18-FEB-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-025-197-107

Query Match 100.0%; Score 1949; DB 14; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDFGALPPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFQSVYWGILTVGSSWG 60  
Db 1 MYDFGALPPEINSARMYAGPGSASLVAAAQMDVSASDLFSAASAFQSVYWGILTVGSSWG 60  
Qy 61 SSAGLWVAASPYVAMSVTAAGAEITTAQOVVAAAAYETAAGLTVPPVIAENRAELMI 120  
Db 61 SSAGLWVAASPYVAMSVTAAGAEITTAQOVVAAAAYETAAGLTVPPVIAENRAELMI 120  
Qy 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGVAAATATATATLLPPEEAPEMTSAG 180  
Db 121 LIATNLGONTPAIAVNEAEYGEEMWAQDAAMFGVAAATATATATLLPPEEAPEMTSAG 180

QY 181 LLEQAAVEEASPTAAANQNMNNVPOALQOAPTOGTTPSSKLGIMKTIVSPHRSPI SN 240  
DB 181 LLEQAAVEEASPTAAANQNMNNVPOALQOAPTOGTTPSSKLGIMKTIVSPHRSPI SN 240  
QY 241 MYSMANNHMSMTNSGVSMNTTSSMLKGFAPAAAQAVOTPAONGVRAMSLGSSIGSSG 300  
DB 241 MYSMANNHMSMTNSGVSMNTTSSMLKGFAPAAAQAVOTPAONGVRAMSLGSSIGSSG 300  
QY 301 LGGGVANLIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGSLPV 360  
DB 301 LGGGVANLIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGSLPV 360  
QY 361 GQMGARAGGSLGSLRVPPRPYPMPHSPAG 391  
DB 361 GQMGARAGGSLGSLRVPPRPYPMPHSPAG 391

## RESULT 13

US-09-072-967-107  
Sequence 107, Application US/09072967  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
APPLICANT: Campos-Neco, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-967-107

Query Match 100.0%; Score 1949; DB 14; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEPINSARVYAGPGSASLVAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60  
DB 1 MVDGALPEPINSARVYAGPGSASLVAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60

QY 61 SSAGLWVAASPVVAMSVTAGQABELTAAQVRVAAAAYETAVGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWVAASPVVAMSVTAGQABELTAAQVRVAAAAYETAVGLTVPPVIAENRAELMI 120  
QY 121 LIATNLIGONTPAIAVNEAEYGEEMWQDAAMGVAAATATATATLLPBEAEEMTSAGG 180  
DB 121 LIATNLIGONTPAIAVNEAEYGEEMWQDAAMGVAAATATATATLLPBEAEEMTSAGG 180  
QY 181 LLEQAAVEEASPTAAANQNMNNVPOALQOAPTOGTTPSSKLGIMKTIVSPHRSPI SN 240  
DB 181 LLEQAAVEEASPTAAANQNMNNVPOALQOAPTOGTTPSSKLGIMKTIVSPHRSPI SN 240  
QY 241 MYSMANNHMSMTNSGVSMNTTSSMLKGFAPAAAQAVOTPAONGVRAMSLGSSIGSSG 300  
DB 241 MYSMANNHMSMTNSGVSMNTTSSMLKGFAPAAAQAVOTPAONGVRAMSLGSSIGSSG 300  
QY 301 LGGGVANLIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGSLPV 360  
DB 301 LGGGVANLIGRAASVGSLSVPOMAAANQAVTPAARALPLTSLTSAERPGOMLGSLPV 360  
QY 361 GQMGARAGGSLGSLRVPPRPYPMPHSPAG 391  
DB 361 GQMGARAGGSLGSLRVPPRPYPMPHSPAG 391

## RESULT 14

US-09-597-796C-8  
Sequence 8, Application US/09597796C  
GENERAL INFORMATION:  
APPLICANT: Skeiky, Yasir  
APPLICANT: Reed, Steven  
APPLICANT: Alderson, Mark  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
FILE REFERENCE: 014058-009050US  
CURRENT APPLICATION NUMBER: US/09/597,796C  
CURRENT FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: WO PCT/US99/07717  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 09/287,849  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 60/158,338  
PRIOR FILING DATE: 1999-10-07  
PRIOR APPLICATION NUMBER: US 60/158,425  
PRIOR FILING DATE: 1999-10-07  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 8  
LENGTH: 391  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: MTB39 (TbH9) protein full-length  
US-09-597-796C-8

Query Match 100.0%; Score 1949; DB 19; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEPINSARVYAGPGSASLVAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60  
DB 1 MVDGALPEPINSARVYAGPGSASLVAAQMDVSASDLFSAASAFOSVVMGLTVGSWIG 60  
QY 61 SSAGLWVAASPVVAMSVTAGQABELTAAQVRVAAAAYETAVGLTVPPVIAENRAELMI 120  
DB 61 SSAGLWVAASPVVAMSVTAGQABELTAAQVRVAAAAYETAVGLTVPPVIAENRAELMI 120  
QY 121 LIATNLIGONTPAIAVNEAEYGEEMWQDAAMGVAAATATATATLLPBEAEEMTSAGG 180  
DB 121 LIATNLIGONTPAIAVNEAEYGEEMWQDAAMGVAAATATATATLLPBEAEEMTSAGG 180

Db 121 LIAITMLGONTPAIAVNEAEYEGEMMAQDAAMFGVAAATATATATLLPFEAPBEMTSAGG 180  
QY 181 LIEQAAAVEASDTPAAANOLMNNVPOLAOPTGTTSSKLGGLMKTVPSPHRSPISN 240  
Db 181 LIEQAAAVEASDTPAAANOLMNNVPOLAOPTGTTSSKLGGLMKTVPSPHRSPISN 240  
QY 241 MWSMANNNHSMSTNGSVSMNTTSSMLKGFAPAAAQAVOTAAQNGVRAMSSIGSSIGSSG 300  
Db 241 MWSMANNNHSMSTNGSVSMNTTSSMLKGFAPAAAQAVOTAAQNGVRAMSSIGSSIGSSG 300  
QY 301 LGGVAAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
Db 301 LGGVAAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
QY 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAAAG 391  
Db 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAAAG 391

RESULT 15

US-09-688-672A-26  
; Sequence 26, Application US/09688672A  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Reed, Steven  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Lodes, Michael L.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009041US  
; CURRENT APPLICATION NUMBER: US/09/688,672A  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/158,338  
; PRIOR FILING DATE: 1999-10-07  
; PRIOR APPLICATION NUMBER: US 60/158,425  
; NUMBER OF SEQ ID NOS: 202  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 391  
; TYPE: PRF  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: TBH9 (McB39A)  
US-09-688-672A-26

Query Match 100.0%; Score 1949; DB 20; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.7e-155;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFGALPEINSARMYAGPGSASLVAAAQMWDSVASDLFSAASAFQSVVMGLTVGSMWIG 60  
Db 1 MVDFGALPEINSARMYAGPGSASLVAAAQMWDSVASDLFSAASAFQSVVMGLTVGSMWIG 60  
QY 61 SSAGLMVAAASPYVAMSTYTAQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
Db 61 SSAGLMVAAASPYVAMSTYTAQAEITAAQVRAAAAYETAYGLTVPPVIAENRAELMI 120  
QY 121 LIATMLGONTPAIAVNEAEYEGEMMAQDAAMFGVAAATATATLLPFEAPBEMTSAGG 180  
Db 121 LIATMLGONTPAIAVNEAEYEGEMMAQDAAMFGVAAATATATLLPFEAPBEMTSAGG 180  
QY 181 LIEQAAAVEASDTPAAANOLMNNVPOLAOPTGTTSSKLGGLMKTVPSPHRSPISN 240  
Db 181 LIEQAAAVEASDTPAAANOLMNNVPOLAOPTGTTSSKLGGLMKTVPSPHRSPISN 240  
QY 241 MWSMANNNHSMSTNGSVSMNTTSSMLKGFAPAAAQAVOTAAQNGVRAMSSIGSSIGSSG 300  
Db 241 MWSMANNNHSMSTNGSVSMNTTSSMLKGFAPAAAQAVOTAAQNGVRAMSSIGSSIGSSG 300  
QY 301 LGGVAAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGOMLGGLPV 360

Db 301 LGGVAAANLGRAASVGSLSVPOMAAANAQVTPAARALPLTSLTSAERPGOMLGGLPV 360  
QY 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAAAG 391  
Db 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAAAG 391

Search completed: November 21, 2003, 17:09:44  
Job time : 174 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 16:56:04 ; Search time 17 Seconds  
(without alignments)  
1081.614 Million cell updates/sec

Title: US-09-724-685-107  
Perfect score: 1949  
Sequence: 1 MVDFGLPPEINSARMYAGP.....SGVLKVPPEPYMHPSPAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1656.5	85.0	396	1 YD61_MYCTU	Q11031 mycobacteri
2	705	36.2	408	1 YS92_MYCTU	Q10813 mycobacteri
3	688.5	35.3	463	1 Y102_MYCTU	Q53951 mycobacteri
4	444	22.8	487	1 Y442_MYCTU	P42611 mycobacteri
5	426.5	21.9	443	1 Y878_MYCTU	Q10540 mycobacteri
6	418	21.4	678	1 YF48_MYCTU	Q10778 mycobacteri
7	364.5	18.7	408	1 SRA_MYCLE	Q10797 mycobacteri
8	324.5	16.6	463	1 Y096_MYCTU	Q10892 mycobacteri
9	324	16.6	434	1 YU8 MYCTU	P31500 mycobacteri
10	321.5	16.5	435	1 YU21_MYCTU	O53268 mycobacteri
11	233.5	12.0	178	1 Y129_MYCTU	O06246 mycobacteri
12	217.5	11.2	176	1 Y125_MYCTU	Q50703 mycobacteri
13	153.5	7.9	860	1 ELS_MOUSE	P54320 mus musculu
14	151.5	7.8	232	1 Y126_MYCTU	O50702 mycobacteri
15	150	7.7	864	1 ELS_RAT	O99372 rattus norv
16	143.5	7.4	730	1 ELS_HUMAN	P15502 homo sapien
17	139.5	7.2	881	1 PRJ3_YEAST	P47033 saccharomyc
18	135	6.9	907	1 A180_HUMAN	O60641 homo sapien
19	133.5	6.8	825	1 ICPO_HSV2H	P28284 herpes simp
20	132	6.8	1120	1 STRF_ECOLI	P76072 escherichia
21	132	6.8	2090	1 N214_HUMAN	P35658 homo sapien
22	131	6.7	836	1 VG26_BPMLS	O05233 mycobacteri
23	131	6.7	1150	1 APMU_PIG	P12021 sus scrofa
24	129	6.6	1783	1 RAA3_CHLRE	O91634 chlamydomon
25	128	6.6	779	1 SRP_DROME	P24856 notocholla
26	127.5	6.5	790	1 ANP_NORCO	P52172 dirosophila
27	126	6.5	1354	1 YAUJ_SCHPO	O10169 schizosacch
28	125.5	6.4	1211	1 BLN2_DROME	Q24523 dirosophila
29	125	6.4	577	1 CS12_HUMAN	P33340 homo sapien
30	124	6.4	2090	1 HFC1_WESAU	P51611 mesocricetu
31	123.5	6.3	677	1 Y136_MYCTU	O50597 mycobacteri
32	123	6.3	432	1 YF10_MYCTU	P71789 mycobacteri
33	123	6.3	1140	1 YM96_YEAST	Q04893 saccharomyc

34	122	6.3	394	1 HYF1_ALCEU	P45805 alcaligenes
35	121.5	6.2	1025	1 SLAP_CAUCR	P35828 caulobacter
36	121.5	6.2	1845	1 Z236_HUMAN	O90136 homo sapien
37	121	6.2	1199	1 P121_RAT	P52591 rattus norv
38	119.5	6.1	635	1 HMLA_DROME	P10105 dirosophila
39	119	6.1	915	1 A180_RAT	O05140 rattus norv
40	118.5	6.1	580	1 EXPR_XANCP	P23314 xanthomona
41	118.5	6.1	2038	1 FSH_DROME	P13709 dirosophila
42	118	6.1	1508	1 BCSC_XANNC	P58938 xanthomona
43	117	6.0	444	1 Y808_CHLRE	O92798 chlamydia p
44	117	6.0	652	1 P1CA_HUMAN	Q13492 homo sapien
45	117	6.0	774	1 STP_LAMB	P03764 bacteriophag

## ALIGNMENTS

RESULT 1  
ID YD61\_MYCTU STANDARD; PRT; 396 AA.  
AC Q11031;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical PPE-family protein RV1361c.  
DE RV1361c OR MT1406 OR MTCY02H10.25C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_Taxid=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; Pubmed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulten J.E., Taylor K., Whitehead S., Barrett B.G.,  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains";  
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Z75555; CAAG9966.1; -;  
CC EMBL: AE007013; AAK45663.1; -;  
CC PIR: H70741; H70741.  
CC TIGR: MT1406; -;  
CC Tuberculist: RV1361c; -;  
CC InterPro: IPR000030; Microdb\_PPE.  
CC Pfam: PF00823; PPE; 1.

KM Hypoetical protein; Complete proteome.  
 FT CONFLICT 158 159 TA -> AT (IN REF. 2).  
 SQ SEQUENCE 396 AA, 40015 MW, 6AFAE0D7B5F66800 CRC64;  
 Query Match 85.0%; Score 1656.5; DB 1; Length 396;  
 Best Local Similarity 85.1%; Pred. No. 4.8e-89;  
 Matches 338; Conservative 18; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPEPINSARMYAGPGSASLVAAQWDSVASDLSAASAFQSVVWGLTVGSWIG 60  
 DB 1 MVDGALPEPINSARMYAGPGSASLVAAQWDSVASDLSAASAFQSVVWGLTVGSWIG 60  
 QY 61 SSAGLWVAASPYVAMWSVTAAGAEITPAQVRAVAAAYETAYGTVPPPIAENRAELMI 120  
 DB 61 SSAGLWVAASPYVAMWSVTAAGAEITPAQVRAVAAAYETAYGTVPPPIAENRAELMI 120  
 QY 121 LIATNLGONTPAIAVNAEAYGEMWADAAAFGYAATAATATATLLPFEAPEMTSAGG 180  
 DB 121 LIATNLGONTPAIAVNAEAYGEMWADAAAFGYAATAATATATLLPFEAPEMTSAGG 180  
 QY 181 LLEQAAVAASTPAANQIMNNVPQALQOAPTOGTTSSKLGKMTVSPHRSPTSN 240  
 DB 181 LLEQAAVAASTPAANQIMNNVPQALQOAPTOGTTSSKLGKMTVSPHRSPTSN 240  
 QY 241 MVSAMNNHSMSTNSGVSMNTLSSMLKFAPAAAAOAVOTPAONGVRAMS---LGSLL 296  
 DB 241 IVSMNNHVSMTNSGVSMNTLSSMLKFAPA---AAQAVETPAQNGVQAMSSLSGQLSSL 299  
 QY 297 GSSGLGGVGAANIGRAASVGSLSVPQAMAAANQAVTPAARALPLTSLTSAERPGQMLG 356  
 DB 300 GSSGLGGVGAANIGRAASVGSLSVPQAMAAANQAVTPAARALPLTSLTSAERPGHMLG 359  
 QY 357 GLPVGMGAAG--GGLSGLVRPPRPVYVMPHSPAAG 391  
 DB 360 GLPLGQLTNSGGFGVSNALRMPPRAYVMPRVPAAG 396

RESULT 2  
 YS92\_MYCTU STANDARD; PRT; 408 AA.  
 ID YS92\_MYCTU STANDARD; PRT; 408 AA.  
 AC Q10813;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypoetical PPE-family protein Rv2892c.  
 GN Rv2892c OR MT2959 OR MTC1274.23C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NC NCBI\_TaxID=1773;  
 RX MEDLINE=9829587; PubMed=6634230;  
 RA Gordon S.T., Brosch R., Parthill J., Garnier T., Churcher C., Harris D.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornby T., Jagers K., Krogh A., McLean A., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares R.,  
 RA Suleron J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gilm M.L., Hatt D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolova M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,  
 RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -I- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z74024; CA98377.1; -;  
 DR EMBL: AB007119; AAK47285.1; -;  
 DR PIR: G70925; G70925.  
 DR TIGR: MT2959; -;  
 DR TubercuList; Rv2892c; -;  
 DR InterPro: IPR000030; Microbac\_PPE.  
 DR Pfam: PF00823; PPE; 1.  
 DR Hypoetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 56 POTENTIAL.  
 SQ SEQUENCE 408 AA, 41469 MW, 3E3D1F20D7827199 CRC64;  
 Query Match 36.2%; Score 705; DB 1; Length 408;  
 Best Local Similarity 41.8%; Pred. No. 4.4e-34;  
 Matches 182; Conservative 47; Mismatches 134; Indels 72; Gaps 14;

QY 2 VDEGALPEPINSARMYAGPGSASLVAAQWDSVASDLSAASAFQSVVWGLTVGSWIGS 61  
 DB 1 MDEGALPEPINSARMYAGPGSASLVAAQWDSVASDLSAASAFQSVVWGLTVGSWIGS 60  
 QY 62 SAGLWVAASPYVAMWSVTAAGAEITPAQVRAVAAAYETAYGTVPPPIAENRAELMI 121  
 DB 61 AASWVAATVPYVAMWSVTAAGAEITPAQVRAVAAAYETAYGTVPPPIAENRAELMI 120  
 QY 122 LIATNLGONTPAIAVNAEAYGEMWADAAAFGYAATAATATATLLPFEAPEMTSAGG 181  
 DB 121 VATTNFGONTPAIAVNAEAYGEMWADAAAFGYAATAATATATLLPFEAPEMTSAGG 179  
 QY 182 LLEQAAA---VEASDTAAANQIMNN---VP--QALQO--LAOPTGTTSSKLGKMTVSPHRSPTSN 240  
 DB 180 AQAALAVTSSTVPLTAAVAPOLLQOLSTSLIPWYSALQOVLAEMLGLTFDNNMTIV 239  
 QY 225 -----GGLMKTVSPHRSPTSNMVSAMNNHSMSTNSGVSMNTLSSMLKFAPAARAAA 275  
 DB 240 RLIGISYFDEGL-----LQFPASLAQQAIPETPGAG--DSGSSVLDSWGPITFA 287  
 QY 276 QAVQTAONGVRAMSGL--GSSLSGS-----SGLGGVGAANIGRAASVGSLS 319  
 DB 288 -----GPRASPVAGGAVGVQTPQPIYVWALDRESIGSVSALGKSSAGSLS 338  
 QY 320 VPOAMAAANQAVTPAARALP--LTSLSAERPGQMLGGLPVGMGAAGGLSGVLR 376  
 DB 339 VPEMDAARARWNPAPMARLPGDDVTALRGTAENA--LIRGFMASAGSTGGGF--VHK 393  
 QY 377 VPPRPVYVMPHSPAAG 391  
 DB 394 YGFRLAVMORPPFAG 408

RESULT 3  
 Y102\_MYCTU STANDARD; PRT; 463 AA.  
 ID Y102\_MYCTU STANDARD; PRT; 463 AA.  
 AC O53951;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypoetical PPE-family protein Rv1802.  
 GN Rv1802 OR MT1851 OR MTV049.24.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

```

OK NCBI_TaxID=1773;
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruter S., Seeger K., Skelton S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, AL022021; CAI17723.1; -
CC EMBL, AB007044; AAK46123.1; -
CC PIR: C70931; C70931.
DR TIGR; MT1851; -
DR TubercuList; Rv1802; -
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KM Hypothetical protein. Complete proteome.
FT CONFLICT 401 401 S -> L (IN REF. 2).
SQ SEQUENCE 463 AA; 46021 MW; EE64828BF09FA551 CRC64;

Query Match 35.3%; Score 688.5; DA 1; Length 463;
Best Local Similarity 42.8%; Pred. No. 4.5e-33;
Matches 174; Conservative 48; Mismatches 136; Indels 49; Gaps 11;

QY 2 VDFGALPEINSGARMYAGFGSASLVAAQMDVSADLFSASAFQSVVWGLTVGSGWGS 61
DB 1 MDFGVLPPEINSGRMVAGFGSGPMLAAAMADGLATELGSTADYGSVSVLT-GVWSGQ 59
QY 62 SAGLVAAASPVYVWMSVTAGQALTAQVRVAAAAYETAYGLTTPPPIAENRAFLML 121
DB 60 SSGTMAAAAPYVAAWMSATPAALAREAAQASAAAYEAFAATYPPVPAANRFLAEL 119
QY 122 INTNLGONTPAIAVEAEYEGEMWADAAAMFGYAATATATATLLPPEADEMTSAGL 181
DB 120 AATNIFGONTGAIIAAEARYAEKMODAAAMGYAGSSVAT-QVTFPAAPPTTAAAGL 178
QY 182 LEQAAAVEASDPTAAANQMLNNVPAQLQOLAOPTGTPSSKLGIMKTVS--PHRSP1- 238
DB 179 ATQGVAAVQAQVAGSAGN-ARSLVSEVLEFLA--TAGTNNKTVASIMNAVTGVVYSSVY 235
QY 239 -----SNVSNANNNHMGNTNGCVMTNTLSMLKGFAPAAAQAQVRAQNGVAA 288
DB 236 NSMLGLGFAESKKVLPANDTVISTIFGWFQFKFPFVPPFNDLI PK----- 283
QY 289 MSLSGLSGLG-----SSGLGSG--GVANAGRAASVGLSVPAQMAAANAQVTPAARALPL 340

```

```

DB 284 -SALGAGLGRSAISSGLGSTAPASAGASQAGSVGMSVPPSWMAATPAIRTVAAVFS 342
QY 341 TSLTS--AERGRPGOML-----GSLPYGGMGARGGSLGVLRY 377
DB 343 TQLQVPPAAISGSLLSOMALASVAGALGGAAPARATGFLVGGGRV 389

RESULT 4
ID Y442_MYCTU STANDARD; PRT; 487 AA.
AC Y442_MYCTU O53727;
DT 01-NOV-1995 (Rel. 32, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein Rv0442c.
GN Rv0442C OR MT0458 OR MT037.06C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Brumann;
RX MEDLINE=87137260; PubMed=3029018;
RA Shinnick T.M.;
RL "The 65-kilodalton antigen of Mycobacterium tuberculosis.",
RL J. Bacteriol. 169:1080-1086(1987).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruter S., Seeger K., Skelton S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, M15467; AAA88235.1; ALT INIT.
DR EMBL, AL021932; CAI17399.1; -
DR EMBL, AB006948; AAK44681.1; -
DR PIR: C70830; C70830.
DR TIGR; MT0458; -
DR TubercuList; Rv0442C; -
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF01469; Pentapeptide_2; 5.

```

DR Pfam; PF00823; PPE; 1.  
 KM Hypothetical protein; Complete proteome.  
 FT CONFLICT 40 E > K (IN REF. 2).  
 FT CONFLICT 96 I -> T (IN REF. 1).  
 FT CONFLICT 211 G -> GNNNG (IN REF. 1).  
 SQ SEQUENCE 487 AA; 47247 MW; 972345B316C8CF CRC64;

Query Match 22.8%; Score 444; DB 1; Length 487;  
 Best Local Similarity 32.1%; Pred. No. 6,3e-19;  
 Matches 135; Conservative 48; Mismatches 180; Indels 58; Gaps 12;

QY 4 FGALPEINSAARYAGSGASLVAAQAQWDSVASDLFSAAAFOSVYWGTLTSSWISSA 63  
 DB 6 FAWLPPEINSAALMPAGSGPLIAAATWAMELABELLASISGSVTSLSGAWLGPSSA 65  
 QY 64 GLMVAASPVYAAWMSVTAGAEELTAQVRAAAAYETAYGLTPPPYIAENRAELMLIA 123  
 DB 66 AAMVAATQYLAWLSTAAQAQEAQAAQAMAATAFEMALAAATQPAVVAANRGLMQLAA 125  
 QY 124 TNLGONTPAIVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAG---- 179  
 DB 126 TNMFGQAPALMDVEAAVEQWMLDVAAAGYHFDASAAVAQAAPWQV--LRNLGIDIG 183  
 QY 180 --GLLEQAAAVEBASDTAAANQMLNNYPAQLQOLAPTOGTSSKLG-----GLW 228  
 DB 184 KNGQINLPGFNTSGNIGNNNIGNNTGSGNTGTGNTSGNTSGNIGLNLGDGNTFG 243  
 QY 229 KTVS-----PHRSPINNMVSMMNNHSMNTSGVSMNTYLSMKLRPAPAAAQAVQ 279  
 DB 244 NTSGNIGFQITGHDQWFGFGFNSGSGN-IGFSGTGTGNVGLFNS-----287  
 QY 280 TAAQNGRAMSSIGSSIGSGGLGQVAAANGRAASVGLSPQAWAANQAVTPAARALP 339  
 DB 288 GSGNIGIGNSGSLNSGIGTSGT--INAGLG--SASGLNT-SFWMGNQNALGSAAGS 340  
 QY 340 LLSLTSAAERGPQGM-----LGGLPVQMGARAG--GGLSGVLR--VPPRPYVPHSPAA 390  
 DB 341 EALVVSAGVATGMSFRAALSGLTASLSTGSLQGLANLVNSGLTNPVAAAPAPV 400  
 QY 391 G 391  
 DB 401 G 401

RESULT 5  
 Y878\_MYCTU STANDARD; PRT; 443 AA.  
 ID Y878\_MYCTU  
 AC Q10540;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical PPE-family protein RV0878C.  
 GN RV0878C OR MT0901 OR MTCY31.06C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NC NCBL\_TaxID=1773;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=8295987; Pubmed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skellern S., Squares S., Squares R.,  
 RA Stalson J.E., Taylor K., Whitehead S., Barrell B.G.,  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RL complete genome sequence.";  
 RN Nature 393:537-544(1998).  
 [2]

RP SEQUENCE FROM N.A.  
 RC SRRAIN=CDC 1551 / Oshkosh;  
 RA Riechmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Desoy R., Dodson R., Gwim M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z73101; CA97385.1; -;  
 DR EMBL; AB006977; AAK45143.1; ALT\_INIT.  
 DR PIR; C70780; C70780.  
 DR TIGR; MT0901; -;  
 DR TubercuList; RV0878C; -;  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR InterPro; IPR002989; Mycobac\_pentapep.  
 DR Pfam; PF01469; Pentapeptide\_2; 4.  
 DR Pfam; PF00823; PPE; 1.  
 KM Hypothetical protein; Transmembrane; Repeat; Complete proteome.  
 FT TRANSMEM 15 35 POTENTIAL.  
 FT TRANSMEM 38 58 POTENTIAL.  
 FT TRANSMEM 59 79 POTENTIAL.  
 FT TRANSMEM 181 201 POTENTIAL.  
 FT DOMAIN 64 73 POLY-ALA.  
 FT DOMAIN 81 115 ALA-RICH.  
 FT DOMAIN 231 270 4 X 10 AA APPROXIMATE REPEATS.  
 SQ SEQUENCE 443 AA; 43592 MW; C58BEC07F0675E2 CRC64;

Query Match 21.9%; Score 426.5; DB 1; Length 443;  
 Best Local Similarity 31.2%; Pred. No. 5.8e-18;  
 Matches 125; Conservative 56; Mismatches 132; Indels 87; Gaps 14;

QY 2 VDFGALPEINSAARYAGSGASLVAAQAQWDSVASDLFSAAAFOSVYWGTLTSSWISSA 58  
 DB 1 MNFMVLPPEVNSARITAGAPAPMLAAVAWQGLAEIGMAASFSLLISGLTAGSGSAW 60  
 QY 59 IGSSAGLMTVAASPVYAAWMSVTAGAEELTAQVRAAAAYETAYGLTPPPYIAENRAEL 118  
 DB 61 QGPAALMAAAMAAAPYLSLMTAAARABEGAAAGKAAAYEADARATAPALVAAARNOL 120  
 QY 119 MILIATNLLGONTPAIVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTS 178  
 DB 121 LSLVLSNLFGQMLPALAATEASVEQLMAQDVAAWVGHGASVSAQLPWPQO----- 173  
 QY 179 GGLLEQAAAVEBASDTAAANQMLNNYPAQLQOLAPTOGTSSKLGMLKTVSPHRSPI 238  
 DB 174 --LLSVLPVVTAPAGAV-----GVPA--LAIPALGV--ENIG-----V 208  
 QY 239 SNV-----SMANNHSMNTSG-----VSMTN-----TLSSMLKGFAPAAAAQAV 278  
 DB 209 GNLFLGIGNTGNNGNNGNTGDTNFGIGNANVLGNNGNNGNNGNNGNNGNNGNNG 268  
 QY 279 QTAQNGVAMSSIGS-SIGSSIGLGGVAA--NLGPAASVGLSVQAWAANQAVTPAA 335  
 DB 269 NTMFGSGNAGFNIGSGNGSGNLGFGNAGDDVTG-----WNSGD-----309  
 QY 336 RALPLTSLTSAERGPQMLGL--PVGQMGARAGGLSG 373  
 DB 310 -----TNTGGFNSGDIINTGTSPTQTGVANSNGNTG 341



```

RESULT 6
YF48 MYCTU STANDARD; PRT; 678 AA.
ID YF48 MYCTU STANDARD; PRT; 678 AA.
AC Q10778;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein Rv1548c.
RV1548C OR MT1599 OR MTCY48.17.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv; PubMed=9634230;
RX MEDLINE=98295987; Garnier T., Churcher C., Harris D.,
RA Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekala F.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry G.E. III, Tekala F.,
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Kirogh A., McLean U., Moutle S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gmin M.L., Hart D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the European Bioinformatics Institute and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z74020; CAA98335.1; -.
DR EMBL; AB007026; AAK45866.1; ALT_INT.
DR PIR; A70762; A70762.
DR TIGR; MT1599; -.
DR TubercuList; Rv1548c; -.
DR InterPro; IPR000030; Microbac_PPE.
DR InterPro; IPR002989; Mycobac_Pentapep.
DR Pfam; PF01469; Pentapeptide_2; 11.
DR HypoThetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 180 200
FT CONFLICT 258 258 D -> G (IN REF. 2).
SQ SEQUENCE 678 AA; 66736 MW; 209F1593D52533A2 CRC64;

```

```

Query Match 21.4%; Score 418; DB 1; Length 678;
Best Local Similarity 21.4%; Pred. No. 2,9e-17;
Matches 106; Conservative 48; Mismatches 140; Indels 38; Gaps 7;

```

62 SAGIMVAAAPVYVMMSTVTAQAEELTAQGVRAAAAVETAYGLTVPPIAENRAELMIT 121

```

RESULT 7
SRA MYCTE STANDARD; PRT; 408 AA.
ID SRA MYCTE STANDARD; PRT; 408 AA.
AC Q07297;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine-rich antigen (25L) (45 kDa protein).
GN SRA OR ML0411 OR MLC383.14.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93239328; PubMed=8478104;
RA Vega-Lopez F., Brooks L.A., Dockrell H.M., de Smet K.A.L.,
RA Thompson J.K., Hussain R., Stoker N.G.;
RT "Sequence and immunological characterization of a serine-rich antigen
RT from Mycobacterium leprae."
RL Infect. Immun. 61:2145-2153(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95020554; PubMed=7934845;
RA Rinke de Wit T.F., Clark-Curtiss J.E., Abebe F., Kolon A.H.J.,
RA Jonson A.A.M., Thole J.E.R.;
RT "A Mycobacterium leprae-specific gene encoding an immunologically
RT recognized 45 kDa protein."
RL Mol. Microbiol. 10:829-838(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Dutfoy S., Feltham T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean U., Moutle S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the European Bioinformatics Institute and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; U00015; AAC43220.1; -
DR EMBL; X68431; CAA48480.1; -
DR EMBL; Z21952; CAA79950.1; -
DR EMBL; Z297179; CAB09938.1; -
DR EMBL; AL583918; CAC29919.1; -
DR PIR; C86960; C86960.
DR PIR; S33522; S33522.
DR PIR; S39872; S39872.
DR Leproma; ML0411; -
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Antigen; Repeat; Complete proteome.
FT DOMAIN 192 196 POLY-SER.
FT REPEAT 209 235 2 X 6 AA REPEATS OF S-V-A-Q-S-E.
FT REPEAT 209 214 1.
FT REPEAT 230 235 2.
FT CONFLICT 132 132 T -> S (IN REF. 2).
FT CONFLICT 189 189 S -> L (IN REF. 2).
FT CONFLICT 191 191 H -> D (IN REF. 2).
FT CONFLICT 292 292 P -> L (IN REF. 2).
SQ SEQUENCE 408 AA; 42466 MW; 5C0C2BED0D6E6A9D8 CRC64;

Query Match 18.7%; Score 364.5; DB 1; Length 408;
Best Local Similarity 26.3%; Pred. No. 2e-14;
Matches 115; Conservative 72; Mismatches 163; Indels 87; Gaps 11;

QY 1 MVDPGALPPEINSARMTYAGPSASLVAAQWDSVASDLPSAAGAFQSVWGLTVGSWIG 60
DB 1 MFDPMTYSPENVNATLMRGSGSTPLKGAHEAMTSLAQLEAAQEVSDTVAVAVPSAFG 60
QY 61 SSAGLWVAASPYVWMSVTAGQELTPAQVVAAYETAYGLTPPYIAENRAELMI 120
DB 61 ETSMDLASRVSTFYVAMLDGNAENAGLJARVLAHVAAYAFEEBAGMVFLLTVLGNIIHTMA 120
QY 121 LIATNLGONTPAIVANAEYGEWADDAAMFGYAAATATATTLTPFEAPMTSAGG 180
DB 121 LKAINMGQVSTYVAALEADYDLMVWVNSTMTTYRDTVLRETKMENFEPAQVLSR-Y 179
QY 181 LLEQAAVEEASDPAANQNLNNVPALQOLAQ----- 213
DB 180 CMRRDVSNSHSSSSDLSYESTIDNLYDSVQSEEHGSDMSQSYVTCSSVQSELCDS 239
QY 214 -----PTG-----GTPPSKLGGLWKTVPSPHSPISNVSNMNMHMTNSGVSMTNT 261
DB 240 PFGTPSGSQSNDLSATSLTQQLGL-----DSIISSASGLTTNS--ISSST 286
QY 262 LSSMLKFPAAAAQVTAQNGVYRAMSSLSGLSSGLGGVYAN-----LGRASVY 316
DB 287 ASSIM-----PIVASQVETELGRSQV-AVEKMIOSISSTAVSDVAASKVAVAGVQAASV 341
QY 317 SLSPVQAAANQAVTPPAARALP--LTSLSAABRGQGMLGSLPVQGMGABAGGSLSGV 374
DB 342 ALRVEPMWATDSQVVAATASHVPSAGSAITTA-----VSGPLEGV 381
QY 375 LRVPPRPYVMPHSPAG 391
DB 382 TQ-PAEEVLTASVAGG 396

RESULT 8
Y096 MYCTU STANDARD; PRT; 463 AA.
ID Y096 MYCTU STANDARD; PRT; 463 AA.
AC 010892;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein Rv0096.
GN Rv0096 OR MTO105 OR MTCY251.15.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1773;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagele K., Krogh A., Mclean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RN Nature 393:537-544(1998).
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Bishai W., Urdarback T., Weidman J., Khouiri H., Gill J., Mikula A.,
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC - SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC - SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; Z74410; CAA98932.1; -
DR PIR; H70750; H70750.
DR TIGR; MTO105; -
DR TubercuList; Rv0096; -
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 276 296 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 419 439 POTENTIAL.
SQ SEQUENCE 463 AA; 46894 MW; 42D9D66A033D0DD8 CRC64;

Query Match 16.6%; Score 324.5; DB 1; Length 463;
Best Local Similarity 27.8%; Pred. No. 4.8e-12;
Matches 112; Conservative 54; Mismatches 144; Indels 93; Gaps 12;

QY 6 ALPPEINSARMTYAGPSASLVAAQWDSVASDLPSAAGAFQSVWGLTVGSITGSAGL 65
DB 2 ALPPEVHSGILSAGCGPSGSLVAAQWDSVASDLPSAAGAFQSVWGLTVGSITGSAGL 61
QY 66 MVAASPYVWMSVTAGQELTPAQVVAAYETAYGLTPPYIAENRAELMIILATN 125
DB 62 YVAAHSPYLAWEQVTAQNGVYRAMSSLSGLSSGLGGVYAN-----LGRASVY 121
QY 126 LIGONTPAIVANAEYGEWADDAAMFGYAAATATATTLTPFEAPMTSAGLLEQA 185
DB 122 FFGINTVPALNADYVRMVLQADTMAYQAVADAATVAVSTQAPPIRAAG----- 176
QY 186 AAVEASDT-----AAANQNLNNVPALQOLAQPTO-----GTPPSKLG----- 224

```

```

Db      177 ----DAADRLDVLSSIGQLIRDI---LDFIANPYKFFEFQFGSPAVTVALVAL 229
QY      225 ----GGLM-----XTVSPHRSPIGNMVMNMMHSMNTNSGVSMTNTLSSMK-----GF 269
Db      230 QLYDFLMYRYVYASGILLDFFTF-----TISALTALASALITHLNLPPAGI 275
QY      270 AAAAAAQAQVTAQAQNGVRAMSSLGSSSLGSSGCGVAAANIGRAASVGSLSVPQAMAAQ 329
Db      276 LPIAAA-----LGPDDQMGANLAVATPATAAVF-----GGSP 308
QY      330 AATPAPRALPLISLTSAERPG---QMGGLFVG---QMGARAG 368
Db      309 PTSPAPAPASNSVGSASAPGISYAVPGLAPPVSSGPKAG 351

RESULT 9
YU18_MYCTU STANDARD; PRT; 434 AA.
ID      YU18_MYCTU
AC      P31500; O53265;
DT      01-JUL-1993 (Rel. 26, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DE      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical PPE-family protein RV3018C.
GN      RV3018C OR MT3098/MT3101 OR MTV012.32C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinomycetales;
OC      Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA      Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence.";
RL      Nature 393:537-544(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CDC 1551 / Oshkosh;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA      Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W.,
RT      "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains.";
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 160-374 FROM N.A.
RC      STRAIN=Isolate 50410;
RA      Packi A.H., Dale J.W.;
RL      Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC      -1- CAUTION: In strain Oshkosh the gene for this protein is
CC      interrupted in position 307 by an IS6110 element.
CC      -1- CAUTION: Was originally (Ref.3) thought to be a dihydrofolate
CC      reductase.
CC      -1- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC      in positions 294; 337 and 355.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@sib.ch).
CC      -----
DR      EMBL: AL021287; CAA16103.1; -.
DR      EMBL: AE007129; AAK47427.1; ALT_SEQ.
DR      EMBL: AE007129; AAK47430.1; ALT_SEQ.
DR      EMBL: X59271; CAA1961.1; ALT_FRAME.
DR      PIR: E70857; E70857.
DR      TIGR: MT3098; -.
DR      TIGR: MT3101; -.
DR      TubercuList; RV3018C; -.
DR      InterPro; IPR00030; Microbac_PPE.
DR      Pfam; PF00823; PPE; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 434 AA; 43029 MW; 41D673C4BD389DD6 CRC64;

Query Match      16.6%; Score 324; DB 1; Length 434;
Best Local Similarity 28.0%; Pred. No. 4,7e-12;
Matches 115; Conservative 56; Mismatches 178; Indels 62; Gaps 12;

QY      6 ALPPEINSAARMYAGPGSASLVAAQWDSVADLPSAAGAFQSVYWGTVGSMIGSSAGI 65
Db      8 ASPPVSHALLSAGPGPSLQAAAAGMSLSAEYAAVNOELSVVAAVAGAGVWGSAEL 67
QY      66 MYAASPYVWMSVTAGQELTPAQVRVAAAAYETAYGLTVPPEVIAENRAELMILIAIN 125
Db      68 FVAAYVPYVAMLVQASDASAAAAGEHEAAAAGVCAALAMPPLPELAANHILTHAVLVATN 127
QY      126 LIGQNTPAIVNABYGEKWAODAAAMGCAATATATTLTLPFBAPMTSAGLLEBA 185
Db      128 PFGINTPIALNEADYVRMWQAAVTMSAYEAVGAAATVATPHTGPAPIVVRG----- 181
QY      186 AVEEASDPDAAN-----OLMNNVPALQQLAQPTGCTPPSSLGLGIMKTVSPH 234
Db      182 --ANESNNAVAATTPPFMEHIVQFLETPAYVQYLSALSELPA--VAWVWFOLEVD 237
QY      235 ---RSPISNMVSMANNHSMNTNSGVSMNTLSSMKGFA----- 270
Db      238 ILGFNITGITTLASNAQLTFEFAINASVAVAGLLYALAGVIDIVEWYIGNLFGVPL 297
QY      271 -----PAAAQAQVTAQAQNGVRAMSSLGSSSLGSSGCGVAAANIGRAASV-GSLSPVA 323
Db      298 GGPLLGLALAAAVPVAGLAGVAGLAAL-PAVGAA--AGAPALVGSVAPVSGGVSPQA 354
QY      324 WAAQAQVTPAPRALPLISLTSAERPGQMGGLFVGQMGARAGGSLGV 374
Db      355 RLVS--AVEPAPASTSVSL--ASDRGAGAL--GF-VGTAGKESVGOQPGKL 398

RESULT 10
YU21_MYCTU STANDARD; PRT; 435 AA.
ID      YU21_MYCTU
AC      O53268; O53269;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical PPE-family protein RV3021C/RV3022C.
DE      RV3021C/RV3022C OR MT3106 OR MTV012.35C/MTV012.36C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinomycetales;
OC      Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA      Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,

```



Db 57 GSSSDLMADAAAGRYLDMITKHSRQILETAVVIDELAVYEETRRKVPPIATIANNREEVH 116  
 QY 120 ILIATNLGONTPAIVANEAEYEGEMMAQDAAMFGVAATATATLTPFEAPEMTSAG 179  
 Db 117 RLIAISNAGVNTPAIAGDAQYQYRAONTIAVMDYOSTAFIAYIPRWOEPQIYGGG 176  
 QY 180 G 180  
 Db 177 G 177

## RESULT 12

Y25 MYCTU STANDARD; PRT; 176 AA.  
 ID Y25 MYCTU STANDARD; PRT; 176 AA.  
 AC 050703;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical PPE-family protein Rv3425.  
 GN Rv3425 OR M7CY78.04C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 CX NCBI\_TaxID=1773;  
 RN 1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagsels K., Kirogh A., Mclean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z77165; CAB01031.1; -;  
 DR PIR; F70738; F70738;  
 DR TubercuList; Rv3425; -;  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KM Hypothetical protein; Complete proteome.  
 SO SEQUENCE 176 AA; 19855 MW; B8CFE2B946B87B0 CRC64;

Query Match 11.2%; Score 217.5; DB 1; Length 176;  
 Best Local Similarity 33.5%; Pred. No. 2.5e-06;  
 Matches 58; Conservative 31; Mismatches 83; Indels 1; Gaps 1;

QY 7 LPPEINSARMYAGPGSASLVAAAOOMDSVASDLFSAASAFOSVYWGTLTGSWIGSSAGTM 66  
 Db 5 TPPEISYINITEGPGADSLFPAAGQLRELIVSVETTHESJEDLELD-ENMKGSSDL 63  
 QY 67 VAAASPVVAMSVTAQGAELTAAQVRAAAVETAVGLTPPVIENRAELMILATNL 126  
 Db 64 ADAVERLYQLWSKSSQLKHAAMWINGLANAAYNDTRKVVPPPEIANRERRRLIASNV 123  
 QY 127 LGONTPAIVANEAEYEGEMMAQDAAMFGVAATATATLTPFEAPEMTSAG 179  
 Db 124 AGVNTPAIADLDQYQYQYRAONVAVMAYVSWTRSLSDLPFRWRPEPQIYRG 176

## RESULT 13

ELS MOUSE STANDARD; PRT; 860 AA.  
 ID ELS MOUSE STANDARD; PRT; 860 AA.  
 AC P54320;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Elastin precursor (Tropoelastin).  
 GN ELN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN 1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Lung;  
 RX MEDLINE=95130069; PubMed=7829060;  
 RA Wyder K.S., Sechler J.L., Boyd C.D., Passmore H.C.;  
 RT "Use of an intron polymorphism to localize the tropoelastin gene to  
 RT mouse chromosome 5 in a region of linkage conservation with human  
 RT chromosome 7.";  
 RL Genomics 23:125-131(1994).  
 CC -1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND  
 CC NUCAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.  
 CC -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER  
 CC INTO AN EXTENSIBLE 3D NETWORK.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.  
 CC -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U08210; AAA80155.1; -;  
 DR PIR; A55721; EAMS.  
 DR MGD; MGI:95317; Eln.  
 DR InterPro; IPR003979; Tropoelastin.  
 DR PRINTS; PR01500; TROPELASTIN.  
 KM Structural protein; Repeat; Signal; Connective tissue.  
 FT SIGNAL 1 27  
 FT CHAIN 28 860  
 FT DISULFID 850 855  
 SO SEQUENCE 860 AA; 71955 MW; 0C0B55AA1ED7F1 CRC64;

Query Match 7.9%; Score 153.5; DB 1; Length 860;  
 Best Local Similarity 24.4%; Pred. No. 0.069;  
 Matches 115; Conservative 34; Mismatches 192; Indels 131; Gaps 17;

QY 5 GALP-----PEINSARMYAGPGSASLVAAAOOMDSVASDLFSAASAFOSVYWGTLTGSW-- 58  
 Db 402 GGIPIGVGCGIGPGIVGPGAVSPAAAKAAKAAK--YGARGVGCIPTVYGAGAGGFG 459  
 QY 59 -IGSSAGLMVAASPVVAMSVTAQGAELTAAQVRAAAVETAVGLTPPVIENRA 116  
 Db 460 YGVGAGAGL--GGASPAAL--AAAKAKAKYG-----AGGAG 491  
 QY 117 ELMILATNLGONTPAIVANEAEYEGEMMAQDAAMFGVAATATATLTPFEAPEM 175  
 Db 492 ALGGIVPGAIVPGAIVPGAIVPGAIVPGAIVPGAIVPGAIVPGAIVPGAIVPGAIV 551  
 QY 176 TSAGCL-----LEQAAVVEASDTAAANQ----- 199  
 Db 552 VGVGIGIVGAVGV 611  
 QY 200 LNNVFPALQQLAQTPTGT-----TPSSKLGLKVTSPHSPISSNMVANN 247  
 Db 612 LGAGVGFAGAGAGVGFAGAGAGVGFAGAGAGVGFAGAGAGVGFAGAGVGFAGAGAG 671



DR EMBL; M86355; AAA42271.1; JOINED.  
DR EMBL; M86363; AAA42271.1; JOINED.  
DR EMBL; M86364; AAA42271.1; JOINED.  
DR EMBL; M86366; AAA42271.1; JOINED.  
DR EMBL; M86371; AAA42271.1; JOINED.  
DR EMBL; M86376; AAA42272.1; -  
DR EMBL; M86373; AAA42272.1; JOINED.  
DR EMBL; M86375; AAA42272.1; JOINED.  
DR PIR; A36106; EART.  
DR InterPro; IPR003979; tropoelastin.  
DR PRINTS; PR01500; TROPOELASTIN.  
KW Structural protein; Connective tissue; Repeat; signal;  
KW Alternative splicing.  
FT SIGNAL 1 1 BY SIMILARITY.  
FT NON TER 1 21 ELASTIN.  
FT CHAIN 22 864 BY SIMILARITY.  
FT DISULFID 854 859 Missing (in isoform 2, isoform 5, isoform  
FT VARSPLIC 263 307 7 and isoform 8).  
FT VARSPLIC 308 308 /FTId=VSP\_004244.  
FT VARSPLIC 809 823 Missing (in isoform 3, isoform 5, isoform  
FT VARSPLIC 809 823 6 and isoform 8).  
FT VARSPLIC 809 823 /FTId=VSP\_004245.  
FT VARSPLIC 809 823 Missing (in isoform 4, isoform 6, isoform  
FT VARSPLIC 809 823 7 and isoform 8).  
FT VARSPLIC 809 823 /FTId=VSP\_004246.  
SQ SEQUENCE 864 AA; 72786 MW; 456894BE09E79FD4 CRC64;

Query Match 7.7%; Score 150; DB 1; Length 864;  
Best Local Similarity 25.9%; Pred. No. 0.11;  
Matches 102; Conservative 31; Mismatches 153; Indels 108; Gaps 18;

QY 5 GARP-----PEINARVAVAGSGASLVAAQMDVSADLFSASAFQSVWGLTVGSW-- 58  
DB 399 GGPVGVGGPGTGGPGTGGPGTGGPGTGGPGTGGPGTGGPGTGGPGTGGPGTGGPG 456  
QY 59 --IGSSAGLMVAASPYVAMWSTAGQAEVTAQVRVAATAETAYGLTVPPVIAENRA 116  
DB 457 YGVGAGAGL-----GGASQAAA-----AAAKAKYAG----- 486  
QY 117 ELMIATNTLLGONTPAIAVNEAEYGEWMAQDAAMFGVAATATATATLLPPEAPEMT 176  
DB 487 -----GAGTLGGLVPG-----AVPGALPGAVPGALPGAVPALP-GAVPGVP 527  
QY 177 SAGGL---EQAAVEEASPTAANQLMNVPOALQLOPTGTPSS-KLGIMKTVS 232  
DB 528 GTGGVPGAGTPAAATAAAAKAAA-----KAGQYGLPGVGGVGGVGLPGVG 579  
QY 233 PHRSPISSNVSMNNHMTNSGVSM-TNTLSSMLKGF-APAAQAQAVOTPAQNGVRAMS 290  
DB 580 PG-----GVTGIGTGTGTGLVPGDLGAGTPPAKAKAATAKAYRAA 624  
QY 291 SLGSSLGSSGLGGV-----AANLGRAASV--GSLVPOAWAANQAVTPARALPLT 341  
DB 625 GLGAGVPGAGVGVGAGVGFAGAGAGVGFAGAGVGFAGAGVGFAGAGVGFAGAGVGF 681  
QY 342 SLTSAERBPQGMGLGVPQMGARAG-GGLSGV 374  
DB 682 -----GGPGGLGGPGGLGGPGGPGGPGGLGCV 708

Search completed: November 21, 2003, 17:04:39  
Job time : 20 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:01:34 ; Search time 41 Seconds  
(without alignments)  
2460.941 Million cell updates/sec

Title: US-09-724-685-107  
Perfect score: 1949  
Sequence: 1 MVDGALPPEINSARMYAGP.....SGVLRYPPRPYVHPSPAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvlnus:\*  
16: sp\_bacteriapi:\*  
17: sp\_archaeapi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1949	100.0	391	16	005298
2	1648	84.6	393	16	Q8VJZ3
3	1583	81.2	393	16	006341
4	779.5	40.0	411	16	053939
5	754	38.7	410	2	Q99C11
6	751	38.5	409	16	053957
7	737.5	37.8	403	16	053956
8	736	37.8	423	16	053950
9	731	37.5	421	16	Q925K0
10	702	36.0	391	16	P96362
11	702	36.0	413	16	Q8VJW0
12	683.5	35.1	464	16	Q8VJW0
13	682.5	35.0	468	16	053958
14	675	34.6	380	16	P95190
15	668.5	34.3	462	16	033310
16	667.5	34.2	385	16	033204

17	666.5	34.2	385	16	Q8VJZ0	Q8VJZ0 mycobacteri
18	635	32.6	363	16	053940	053940 mycobacteri
19	633.5	32.5	365	16	086373	086373 mycobacteri
20	629.5	32.3	405	16	Q8VJW5	Q8VJW5 mycobacteri
21	618.5	31.7	397	2	Q9AGF0	Q9AGF0 mycobacteri
22	615.5	31.6	402	16	033312	033312 mycobacteri
23	610	30.3	443	16	Q8VKL9	Q8VKL9 mycobacteri
24	603	30.9	426	16	005907	005907 mycobacteri
25	597.5	30.7	391	16	005798	005798 mycobacteri
26	589	30.2	406	16	P71869	P71869 mycobacteri
27	588	30.2	394	16	033205	033205 mycobacteri
28	575.5	29.5	391	16	P95247	P95247 mycobacteri
29	575.5	29.5	415	16	Q8VJK7	Q8VJK7 mycobacteri
30	495.5	25.4	3275	16	Q8VKM3	Q8VKM3 mycobacteri
31	495.5	25.4	3300	16	006304	006304 mycobacteri
32	473	24.3	180	16	053722	053722 mycobacteri
33	452.5	23.0	3716	16	050379	050379 mycobacteri
34	448.5	23.0	580	16	006206	006206 mycobacteri
35	444	22.8	346	16	050424	050424 mycobacteri
36	443	22.7	963	16	007232	007232 mycobacteri
37	443	22.7	3186	16	Q8VKM2	Q8VKM2 mycobacteri
38	439.5	22.6	655	16	053949	053949 mycobacteri
39	429.5	22.0	479	16	P71862	P71862 mycobacteri
40	429.5	22.0	686	16	Q8VJ65	Q8VJ65 mycobacteri
41	428.5	22.0	1053	16	065937	065937 mycobacteri
42	428.5	22.0	1105	16	Q8VJX6	Q8VJX6 mycobacteri
43	423.5	21.7	2523	16	053393	053393 mycobacteri
44	422.5	21.7	615	16	P95249	P95249 mycobacteri
45	418	21.4	3157	16	050378	050378 mycobacteri

## ALIGNMENTS

RESULT 1  
ID 005298 PRELIMINARY; PRT; 391 AA.  
AC 005298;  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DT Hypothetical 39.2 kDa protein (PPE family protein).  
GN RV1196 OR MTC1364.08 OR MT1234.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37Rv;  
RX MEDLINE=9825987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.,  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleitechmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Uitterlind T., Weidman J., Knouri H., Gill J., Mikila A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z93777; CAB07839.1; -  
 DR EMBL: AE007000; AAK45491.1; -  
 DR TIGR; MT1234; -  
 DR TubercList; Kvi196; -  
 DR InterPro; IPR000030; Microbac\_ppp.  
 DR Pfam; PF00823; PPE; 1.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 391 AA; 39158 MW; E40936B3ABDC0F8 CRC64;

Query Match 100.0%; Score 1949; DB 16; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-101;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASLVAAQWMSVSDLSAASAFQSVWGLTVSGWIG 60  
 DB 1 MVDGALPEINSARMYAGPGSASLVAAQWMSVSDLSAASAFQSVWGLTVSGWIG 60  
 QY 61 SSAGLWVAASPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPIAENRAELMI 120  
 DB 61 SSAGLWVAASPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPIAENRAELMI 120  
 QY 121 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLLPFEAPEMTSAGG 180  
 DB 121 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLLPFEAPEMTSAGG 180  
 QY 181 LLEQAAVEBASDPTAANQLMNNVPOALQOAPTOGTPSSKLGIMKTVPSPHRSPISN 240  
 DB 181 LLEQAAVEBASDPTAANQLMNNVPOALQOAPTOGTPSSKLGIMKTVPSPHRSPISN 240  
 QY 241 MVSANNNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSLSGSG 300  
 DB 241 MVSANNNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSLSGSG 300  
 QY 301 LGGVAAALGPAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
 DB 301 LGGVAAALGPAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
 QY 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAG 391  
 DB 361 GOMGARAGGGLSGVLRVPPRPVMPHSPAG 391

RESULT 2  
 ID Q8V123 PRELIMINARY; PRT; 393 AA.  
 AC Q8V123;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE PPE family protein.  
 GN MT3582.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleisichmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey B.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolova M.D., Salzberg S.L.,  
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE007161; AAK47941.1; -  
 DR TIGR; MT3582; -  
 DR InterPro; IPR000030; Microbac\_ppp.  
 DR InterPro; IPR000508; SigPase.  
 DR Pfam; PF00823; PPE; 1.  
 DR PROSITE; PS00501; SPASE I 1; 1.  
 SQ SEQUENCE 393 AA; 39668 MW; 86F0B6779885511 CRC64;

Query Match 84.6%; Score 1648; DB 16; Length 393;  
 Best Local Similarity 84.8%; Pred. No. 4.3e-84;  
 Matches 335; Conservative 20; Mismatches 34; Indels 6; Gaps 3;

QY 1 MVDGALPEINSARMYAGPGSASLVAAQWMSVSDLSAASAFQSVWGLTVSGWIG 60  
 DB 1 MVDGALPEINSARMYAGPGSASLVAAQWMSVSDLSAASAFQSVWGLTVSGWIG 60  
 QY 61 SSAGLWVAASPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPIAENRAELMI 120  
 DB 61 SSAGLWVAASPYVAMSVTAGAELTAAQVRAAAAYETAYGLTVPPIAENRAELMI 120  
 QY 121 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLLPFEAPEMTSAGG 180  
 DB 121 LIATNLGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLLPFEAPEMTSAGG 180  
 QY 181 LLEQAAVEBASDPTAANQLMNNVPOALQOAPTOGTPSSKLGIMKTVPSPHRSPISN 240  
 DB 181 LLEQAAVEBASDPTAANQLMNNVPOALQOAPTOGTPSSKLGIMKTVPSPHRSPISN 240  
 QY 241 MVSANNNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSLSGSG 300  
 DB 241 MVSANNNHSMNTSGVSMNTTSSMLKGFAPAAAQAVOTAQNGVRAMSSIGSSLSGSG 300  
 QY 297 GSSGLGGVAAALGPAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLG 356  
 DB 300 GSSGLGGVAAALGPAASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERGPQMLG 359  
 QY 357 GLPVGOMGARAGGGLSGVLRVPPRPVMPHSPAG 391  
 DB 360 GLPLGH-SVWAGGINNALRVPARAIPRTPAAG 393

RESULT 3  
 ID 006341 PRELIMINARY; PRT; 393 AA.  
 AC 006341;  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)  
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Hypothetical protein RV3478.  
 GN RV3478 OR MTCY13E12.31.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=9825987; PubMed=9634230;  
 RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. II, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares K.,  
 RA Sulston J.E., Taylor K., Whitehead S., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RT Nature 393:537-544(1998).  
 DL Nature 393:537-544(1998).  
 DR EMBL: Z95390; CAB08702.1; -  
 DR TubercList; RV3478; -  
 DR InterPro; IPR000030; Microbac\_ppp.  
 DR Pfam; PF00823; PPE; 1.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 393 AA; 39413 MW; AF4C20C95DME7D04 CRC64;

Query Match 81.2%; Score 1583; DB 16; Length 393;  
 Best Local Similarity 81.5%; Pred. No. 1.8e-80;  
 Matches 322; Conservative 20; Mismatches 47; Indels 6; Gaps 3;

```

QY 1 MDPFGALPPEINSARMYAGPGSASLVAAAQMDVSADLFSASAASAFQSVVMGLTVGSMWG 60
DB 1 MDPFGALPPEINSARMYAGPGSASLVAAAQMDVSADLFSASAASAFQSVVMGLTVGSMWG 60
QY 61 SSAGLMVAASPYVAMSVTAGAELTAAQVRAAAAAYETAYGLTPPPVIAENRAELMT 120
DB 61 SSAGLMVAASPYVAMSVTAGAELTAAQVRAAAAAYETAYGLTPPPVIAENRAELMT 120
QY 121 LATNLGONTPAIAVNEAEYEMMAQODAAAMFGYAAATATATATLLPPEEAPEMTSAGL 180
DB 121 LATNLGONTPAIAVNEAEYEMMAQODAAAMFGYAAATATATATLLPPEEAPEMTSAGL 180
QY 181 LLEQAAVAEASDTAAANQLMNNVPOALQOLAQPTGTTTSSKLGIMKTVPSPHSRISN 240
DB 181 LLEQAAVAEASDTAAANQLMNNVPOALQOLAQPTGTTTSSKLGIMKTVPSPHSRISN 240
QY 241 MYSMAANNHSMSTNGVSMNTTSSMLKGFAPPAALAAQVOTAAQNGVRAMS----IGSSL 296
DB 241 MYSMAANNHSMSTNGVSMNTTSSMLKGFAPPAALAAQVOTAAQNGVRAMS----IGSSL 296
QY 297 GSSGGLGCGVAAANLGRASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERGPOMLG 356
DB 300 GSSGGLGCGVAAANLGRASVGSLSVPOAMAAANOAVTPAARALPLTSLTSAERGPOMLG 359
QY 357 GLPVGQMGARAGGGLSGVLAVPVPVPMHSPAPAG 391
DB 360 GLPLGH-SYNAGSGINNALRVPARAVAIPTPAAG 393

```

## RESULT 4

053939 PRELIMINARY; PRT; 411 AA.

```

AC 053939;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE PPE-family protein.
GN RV1789 OR MT1838 OR MTW049.11.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Biglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLaren J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Pleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey B.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Uitterback T., Weidman J., Kouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL020021; CA17711.1; ALT_INT.
DR EMBL; AE070043; AAK46108.1; -.
DR TIGR; MT1838; -.
DR TubercuList; RV1789; -.
DR InterPro; IPR000030; Microbac_PPE.

```

```

DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 411 AA; 40557 MW; 03C9B5E0590B7DA CRC64;

```

Query Match 40.0%; Score 779.5; DB 16; Length 411;  
 Best Local Similarity 43.7%; Pred. No. 7,6e-36;  
 Matches 179; Conservative 65; Mismatches 129; Indels 37; Gaps 10;

```

QY 2 VDFGALPPEINSARMYAGPGSASLVAAAQMDVSADLFSASAASAFQSVVMGLTVGSMWG 61
DB 19 MDPFGALPPEINSARMYAGPGSASLVAAAQMDVSADLFSASAASAFQSVVMGLTVGSMWG 78
QY 62 SSAGLMVAASPYVAMSVTAGAELTAAQVRAAAAAYETAYGLTPPPVIAENRAELMT 121
DB 79 SSAGLMVAASPYVAMSVTAGAELTAAQVRAAAAAYETAYGLTPPPVIAENRAELMT 138
QY 122 LATNLGONTPAIAVNEAEYEMMAQODAAAMFGYAAATATATATLLPPEEAPEMTSAGL 181
DB 139 LATNLGONTSAIAAAEAYGEMMAQODSAAMYAVAGSSASASA-VTFSTPQIANPTAQ 197
QY 182 LLEQAAVAEASDTAA--NQLMNNVPOALQOLAQPTGTTTSSKLGIMK----- 229
DB 198 GTQAAAVATPAGTAOSTLTETMTGLPNAQSLTSPLOS--NGPLSWIQLIFGPNPPT 256
QY 230 -----TVSPHSRISNMTNSGVSMNTTSSMLKGFAPPAALAAQVOTAA 282
DB 257 SISALLTDLQPPASFPYNTGELPYFSIGMNNFISAKTL-GLISAPAAVA-----AA 310
QY 283 QNGVRAMSSLGSSGLGCGVAAANLGRASVGSLSVPOAMAA--ANOAVTPAARALPLT 341
DB 311 GDAAGLPLGAGMLG-----GGVAAAGLGNAAASVGLSVPVMSGLPVSGAAPLPVS 366
QY 342 SLTSAERGPOMLGGLPVGQMGARAGGGLSGVLAVPVPVPMHSPAPAG 391
DB 367 TVSAAPPEAAGSLGLGLPL---AGAGGAGAGP-RYGRFPVTWABPPFAG 411

```

## RESULT 5

099011 PRELIMINARY; PRT; 410 AA.

```

AC 099011;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE RV1808-like protein.
GN MYC1808 OR OVI1808.
OS Mycobacterium microti.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1806;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=MYC 94-2272, and OV254;
RA Sivodon V., Heym B., Mazancourt P., Gailhard J.-L.;
RT "PPE RV1808 orthologue of Mycobacterium microti.";
RL Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335180; AAK20894.1; -.
DR EMBL; AF335179; AAK20893.1; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
SQ SEQUENCE 410 AA; 40044 MW; 3B9DA3174655A5EA CRC64;

```

Query Match 38.7%; Score 754; DB 2; Length 410;  
 Best Local Similarity 44.5%; Pred. No. 2e-34;  
 Matches 177; Conservative 51; Mismatches 140; Indels 30; Gaps 10;

```

QY 1 MDPFGALPPEINSARMYAGPGSASLVAAAQMDVSADLFSASAASAFQSVVMGLTVGSMWG 60
DB 1 MDPFGALPPEINSARMYAGPGSASLVAAAQMDVSADLFSASAASAFQSVVMGLTVGSMWG 60
QY 61 SSAGLMVAASPYVAMSVTAGAELTAAQVRAAAAAYETAYGLTPPPVIAENRAELMT 120
DB 61 PSSITMAAIVAPYVAMSVTAGAELTAAQVRAAAAAYETAYGLTPPPVIAENRAELMT 120

```

OY	121	IATLTLLCONFPAIVNNEVEYGEMMAODAAAFGVAAAATATATLTPPEEAEMTSAG	180
Dd	121	LVAITIPQONTPALAATEAHVAMMAODAAAVGYVGSSATA-SQLAPSEPPQTINPSA	179
OY	181	LLEOAAAEASDTAA-----NOLMNNVPALCOOLAOPFGTTPSPKGLMKTVSP	233
Dd	180	TAAOSAVVAQAAGAAASSDITAOLOSILPSTLOSL--TTATATSAAG--WDTV--	233
OY	234	HRSPISNVSMNMNHMSMTNSGV---SMTNTLSMLKGFAPAALAAQAVCTAACNGVRAMS	230
Dd	234	-LGSTTTTLANTGPYSIIIGLGAIPGGMWLTFGQIL-GLAQNAPRYALLGGPAAAAIALS	231
OY	291	SLGSSLGS----SGLGGVAAANLGRASVGSISYPQAMAANAQAVTPAAPALPLTSLTS	345
Dd	292	PLAPLRGGIADITELGGCGATGGIAIRAIYVGSISYPQGAAEAAAPVMRAVASVLPGTGAAP	351
OY	346	A-AERGPQMGLGLPYQMGARAGAGGL-----SGULRV	377
Dd	352	ALAAPAAGALFGEMLISSLAGRLACTAVRSGGAARV	389

	RESULT	6
	053957	
ID	PRELIMINARY;	PRT; 409 AA.
AC	053957;	
DT	01-JUN-1998 (TREMBLrel. 06, Created)	
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)	
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)	
DE	PPE-family protein.	
CN	Rv1808 OR MT1856.1 OR MTV049.30.	
OS	Mycobacterium tuberculosis.	
OC	Bacteria; Actinobacteria; Actinomycetales;	
CC	Corynebacterales; Mycobacteriaceae; Mycobacterium.	
OX	NCHI_Taxid=1773;	
FN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RC	MEBLIN=96295987; PubMed=9634230;	
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,	
RA	Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,	
RA	Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,	
RA	Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,	
RA	Hornsby T., Jagels K., Krogh A., McLean J., Mole S., Murphy L.,	
RA	Oliver S., Osborne J., Quail M.A., Rajadream M.A., Rogers J.,	
RA	Rutter S., Seeger K., Skellon S., Squares S., Squares R.,	
RA	Sultson J.E., Taylor K., Whitehead S., Barrall B.G.;	
RT	"Deciphering the biology of Mycobacterium tuberculosis from the	
RT	complete genome sequence."	
RL	Nature 393:537-544 (1998).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CDC 1551 / Oshkosh;	
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,	
RA	Peterson J., Deboy R., Dodson R., Gwinn M., Hatt D., Hickey E.,	
RA	Kojanay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,	
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,	
RA	Bisbal W.;	
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and	
RT	laboratory strains."	
RL	Submitted (APR-2001) to the EMBL/genbank/DDBJ databases.	
DR	EMBL; AL022021; CAI17729.1;	
DR	EMBL; AE007044; AAA6129.1; ALT_INIT.	
DR	TIGR; MT1856.1;	
DR	TubercuList; Rv1808;	
DR	InterPro; IPR000030; Microbac_PPE.	
DR	Pfam; PF00823; PPE; 1.	
KW	Complete proteome.	
SQ	SEQUENCE 409 AA; 39917 MW; 1E1520BACF36379 CRC64;	

```
Qy      2 VDEGALPPINSAARMYAGGSASLIVAAQQMDVSADLFSSAASFQSVVGLTGVSGIGS   61
Dd      1 MDGALPPEINSGRYAGCGSPILAAAAAADLAAELIYSMAASYGTIEGLTYAAPMGP    60
Qy      62 SAGLMVAASPVMNMSVTAGAEILTAAQEVAAAAYETAYGLTPVPPIYAENRBEIMIL   121
Dd      61 SSTITMAAAVAAPVAMWISTYAGAEQAGAQAIAAGVETAPPAATVPPIVEANALLMSL   120
Qy      122 IANTLLGONTPLAVNEAEYEGEMNAODAAAFGVAATAATATTATLLPFEEAPEMTSAGL   169
Dd      121 VANINGQNTPLAATEAHYAEEMWAQDAAMYGAGSSATA-SQLAPPSEPPQTINTSAT   179
Qy      182 LEOAAAVEASDTAAA-----NQLMNNVPOAQOOLAOPTQSTPSSKLGMLKTIVSPH   234
Dd      180 AAGSAVVAQAAGAAASSDITQAQSILSLSTLSLA--TTATATSASAG-WDIY---   232
Qy      235 RSPDISNMVSMANNHMSMTNSGVS----MTNTLSMKLGFAPPAALAAQAVOTAAQNVRAM   289
Dd      233 LGSTITTLNLNGPSYIIIGAIPEGMWLTFGQILGLAQNPBGVALLGGPRACGALSPL   292
Qy      290 SSU-GSSLGS-SGLGGGVANAIGRAASIGSISVPRAMAANAQVTPAARALPLTSLSLA-   346
Dd      293 APIRGYIDITPLPGGATGATIAIRIIVGSGISLVPGMAEAPVNRVAASVLPGTAAPAL   352
Qy      347 AERPGOMLGLPVGMGABAGGCL-----SGVLRV   377
Dd      353 AAAPGALFGEHALSLAGRALAGTVAASSGAARV   388
```

RESULT 7	053956	PRELIMINARY;	PRT;	403 AA.
ID	053956			
AC	053956			
DT	01-JUN-1998 (TrEMBLrel. 06, Created)			
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE	PPE-family protein.			
DN	RV1807 OR MT1856 OR MTW049.29.			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H37Rv;			
RX	MEDLINE=968259587; PubMed=9634230;			
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,			
RA	Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III, Tekala F.,			
RA	Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA	Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,			
RA	Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,			
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,			
RA	Rutter S., Seeger K., Skelton S., Squares S., Squares R.,			
RA	Sulston J.E., Taylor K., Whitehead S., Barrrell B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence.";			
RL	Nature 393:537-544(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CDC 1551 / Oshkosh;			
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,			
RA	Petersen J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E.,			
RA	Kolonyak J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,			
RA	Decher A., Uetebach T., Weidman J., Khouri H., Gill J., Mikula A.,			
RA	Bisbal W.;			
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and			
RT	laboratory strains.";			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL020221; CAA17728.1; -			
DR	EMBL; AE007044; AAK46128.1; -			
DR	TIGR; MT1856; -			
DR	Tuberculist; RV1807; -			

DR InterPro; IPRO00030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome.  
 FT CONFLICT 227 F -> S (IN REF. 2).  
 FT CONFLICT 238 V -> L (IN REF. 2).  
 SQ SEQUENCE 403 AA; 39243 MW; DCEI1880FD1SCBFE CRC64;

Query Match 37.8%; Score 737.5; DB 16; Length 403;  
 Best Local Similarity 42.6%; Pred. No. 1.6e-33;  
 Matches 181; Conservative 60; Mismatches 123; Indels 61; Gaps 12;

QY 2 VDFGALPPEINSGRMVYTGPGPMLAAATVMDGLAVELHATAGVSELSALT-GAMSGP 59  
 DB 1 MDFGALPPEINSGRMVYTGPGPMLAAATVMDGLAVELHATAGVSELSALT-GAMSGP 59  
 QY 62 SAGLWVAASPYVWMSATVAGQAEELTAQVRAVAAAAYETAYGLTTPPVYIAERAEML 121  
 DB 65 ASASMTAAAPVAVWMSATVAVRAEQAGQAEAAAEAFATVPPIEARRAQLMAL 124  
 QY 122 IATNLLGONTPLAVNEAEYGEWMAODAAAMFGYAATATATATLLPFEAPEMTSAGGL 181  
 DB 125 IATNVLGQAPALPAITAEQYAEWMSQDAAAMYGTAQSAAT-QLTFTEPVQTTVAAGL 183  
 QY 182 LEQAAAEEASDPTAA-----NQLMNNVPOALQOLAQPTGT-----TPSS 222  
 DB 184 AAQSAIAIATGASAGAQTTLISQLIAIPSVLQGLSSSTPAATFASGSGLGIVSGSS 243  
 QY 223 KLGLMKTVSPHRSPTISNMVSMANNHMSMTNSGVSM-TNTLS--SMKGFAPAAAQAV 278  
 DB 244 WDKMKALLDPN-----SNFMNTTASGQLFSPNTIAEFLDLGLGVAAAADAGDV 293  
 QY 279 QTAQNGVRAMSGSLGSS-----GSSGLGGGVANIGRAASVGLSPQAMAAQAQAVP 333  
 DB 294 LGEATSG-----GLGALVAPLQSGAGLGTVAAGLGNAAVTGTLSPSMTAAFLASG 348  
 QY 334 AARAL--PLTSLTSAERGPQMLGSLPYGOMGABAGGSLGVLVPP-----RPVMPH 386  
 DB 349 LGSALGCTPMVAPPAVAG---MPGMPGCTMGGGGPG-----RAVPQYGRPFVAVR 398  
 QY 387 SPAG 391  
 DB 399 PPAG 403

RESULT 8  
 ID 053950 PRELIMINARY; PRT; 423 AA.  
 AC 053950;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE PPE-family protein (PPE family protein).  
 GN RVI1801 OR MTW049.23 OR MT1850.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornby T., Jagels K., Kitchell A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence.";  
 RL Nature 393:537-544 (1998).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Uitterback T., Weidman J., Khouri H., Gali J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 laboratory strains.";  
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL022021; CAI17722.1; -  
 DR EMBL; AE07044; AAK46122.1; -  
 DR TIGR; MT1850; -  
 DR Tuberculist; RVI1801; -  
 DR InterPro; IPRO00030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome.

Query Match 37.8%; Score 736; DB 16; Length 423;  
 Best Local Similarity 41.8%; Pred. No. 2.1e-33;  
 Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;

QY 2 VDFGALPPEINSGRMVYTGPGPMLAAATVMDGLAVELHATAGVSELSALT-GAMSGP 59  
 DB 1 MDFGALPPEINSGRMVYTGPGPMLAAATVMDGLAVELHATAGVSELSALT-GAMSGP 59  
 QY 62 SAGLWVAASPYVWMSATVAGQAEELTAQVRAVAAAAYETAYGLTTPPVYIAERAEML 121  
 DB 60 SSTSWASAAAPVAVWMSATVAVRAEELAGQARLAAIAYEAFATVPPIAARRAQLMAL 119  
 QY 122 IATNLLGONTPLAVNEAEYGEWMAODAAAMFGYAATATATATLLPFEAPEMTSAGGL 181  
 DB 120 IATNLFQQTPLAIPMTAEQYAEWMAODAAAMYGTAQSAAT-SRMTAFTEPQTTHGOL 178  
 QY 182 LEQAAAEEASDPTAAAN-----QLMNNVPOALQOLAQPT-----QGTTP-----S 221  
 DB 179 GAQSSAVAGTAATATTAAGNQLQSAFPLQLLSAVPALQGLAFTASQASATPQWTDGNS 238  
 QY 222 SKLGLMKTVSPHRSPTISNMVSMANNHMSMTNSGVSM-TNTLS-SMKGFAPAAAQAVQTA 281  
 DB 239 TELGG--AVTGPYTP-----GVLPSGVPYLLGIGSVL-----V 271  
 QY 282 AONGVRAMSGSLG-----SLGSSGLGGG--VAANLGRAASVGSLS 319  
 DB 272 TONGQVSALDKIGKPIITGALAPLAEPLHTPTLIGSEGLGGSVAGIGRGLVKLS 331  
 QY 320 VQDAMAAQAQAVTPAARALPLTSLTS---AERGPQMLGSLPYGOMGABAGGSLGVL 376  
 DB 332 VQCGMTVAAPETLPPAALQATRLAAAPRTAATDGAALGMAISGLAGRAAAGSTG--- 388  
 QY 377 VPPRPVMPHSPAG 391  
 DB 389 ---HPGSAADAVG 400

RESULT 9  
 ID 0925KO PRELIMINARY; PRT; 421 AA.  
 AC 0925KO;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Putative PPE protein (PPE-family protein).  
 GN MLI182 OR MLCB1701.08C.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1769;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;

RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagsels K., Lacroix C., Mclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simen S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrall B.G.,  
 RT "Massive gene decay in the leprosy bacillus";  
 RL Nature 409:1007-1011(2001).  
 DR EMBL; AL049191; CAB39148.1; -  
 DR EMBL; AL583921; CAC31563.1; -  
 DR Lepioma; ML1182; -  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 421 AA; 43119 MW; 680FDF5FCB65B4A7 CRC64;

Query Match 37.5%; Score 731; DB 16; Length 421;  
 Best Local Similarity 41.1%; Pred. No. 3.9e-33;  
 Matches 174; Conservative 55; Mismatches 160; Indels 34; Gaps 8;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQMDVSASDLFSASAFQSVYMGTLVGSWIG 60  
 DB 1 MPDFAALPPEINSTRMYLGPSSPILTVAAAWVLAKELTAAQGLQSAVEAL-LTFEEG 59  
 QY 61 SSAGLWAAASPYVAMSVTAAGAEITPAQVRVAAAAYETAYGLTVPPIAENBARLMI 120  
 DB 60 ESDAAALERTVPEKMTLQNAASHETATQTVAAAYETATFTVPLVFNRAQACL 119  
 QY 121 LIATNLIGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLTFEEADEMTSAGC 180  
 DB 120 LINSNIIGONSTALAEKEAYTETEMWIDDAAMTSYQASVLEAGATKATFAAPLGVNEVG 179  
 QY 181 L-----LEQAAVEASDTPAANQLMNN-----VQALQQ-----LAQP 214  
 DB 180 LAQEVVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEE 239  
 QY 215 TQGTTPSSKLGK--LWKTVPHRSPISNMVSMANNHSMNTNSGVSMNTLSSMLKGRAPA 272  
 DB 240 QTAVPDSSAAAPQLWGGFQOHLSPLINDTSLMINNHAGMANAGISLWNGMSAKKSLAP- 298  
 QY 273 AAAQVQTAQNGVRAMSSIGSSISGSGLVANIGRAASVGSLSVPOMAAANOAVT 332  
 DB 299 TTTKAAASAKKAGSAVQSTGRGLTSSSGGHVTAQGRASISGLAEVPTWTASQPTV 358  
 QY 333 PAAPALPLTSLTSAERPGQML--GGLPVQOM--GARAGGGLSGVLRVPPRYVMPHSP 368  
 DB 359 AATRAALSPARVAATESESAPLLGGGLPMAPMVPGGSGGTGVNTALRLQPAFVMPRNP 418  
 QY 389 AAG 391  
 DB 419 AAG 421

RESULT 10  
 P96362 PRELIMINARY; PRT; 391 AA.  
 AC P96362;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical 38.1 kDa protein (PPE family protein).  
 GN R1039C OR MTCY10G2.10 OR MT1068.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagsels K., Krogh A., Mclean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton J., Squares S., Squares R.,  
 RA Sultson J.E., Taylor K., Whitehead S., Barrall B.G.,  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence";  
 RL Nature 393:537-544(1998).

RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Knouri H., Gill J., Mikula A.,  
 RA Bishai W.,  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z92539; CAB06873.1; -  
 DR EMBL; AE006989; AAK45319.1; -  
 DR TIGR; MT1068; -  
 DR Tuberculisc; Rv1039c; -  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 391 AA; 38081 MW; AFD3EAFB195C4F CRC64;

Query Match 36.0%; Score 702; DB 16; Length 391;  
 Best Local Similarity 42.3%; Pred. No. 1.4e-31;  
 Matches 166; Conservative 62; Mismatches 150; Indels 14; Gaps 7;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQMDVSASDLFSASAFQSVYMGTLVGSWIGS 61  
 DB 1 MDPGALPPEINSARMYAGAGGPMMAAGAAWNGLAELGTAAASYEVITRLTTESWMPG 60  
 QY 62 SAGLWAAASPYVAMSVTAAGAEITPAQVRVAAAAYETAYGLTVPPIAENBARLMI 121  
 DB 61 ASIAWAAAPQPYAMLTLYTAAEAHAGSQAMASAAAYEAIVMTVPEVVAARALLAAL 120  
 QY 122 IATNLIGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLTFEEADEMTSAGC 181  
 DB 121 VAINVIGINTPAIMAEALYEMMODALAMGYAASG--AGMLQPLSPSGTTTPGGL 179  
 QY 182 LEQAAVEASDTPAANQ-----LNNVPOALQOLAQPTQGTTPSSKLGMLKTVSPHRS 236  
 DB 180 AAQSAIVGSAATAAANQSVADLISLPNAVSGLASPVSVDLSGLGIADIDALLA 239  
 QY 237 P--ISNMVSMANNHMS--MTNSGVSMNTLSSMLKGRAPAANAQVQTAQNGVRAMSSIG 293  
 DB 240 TTPVANIINSVAVTAAVYNAALPTATFLANALNSGAPVATGAIEAEG--AASAAA 296  
 QY 294 SSLSGSLGCVANIGRAASVGSLSVPOMAAANOAVTPAARALPLTSLTSAERPGQ 353  
 DB 297 AGLADSVTPAGLCAISGEATIVGRSLVPAAMGTAARATTAAGTALGSGMTVAABEA--GP 355  
 QY 354 MUGLFPVQMGARAGGSLGVLVPPRYVMP 385  
 DB 356 VTGMMMP--GMAAAGKGAAGAPRYGKPTVMP 386

RESULT 11  
 O06386 PRELIMINARY; PRT; 413 AA.  
 ID O06386;  
 AC O06386;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical 40.7 kDa protein (PPE family protein).  
 GN RV3621C OR MTCY15C10.31 OR MTCY07H7B.01 OR MT3723.

OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriinae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxId=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornby T., Jagels K., Krogg A., McLean J., Moul S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares R.,  
RA Snelson J.E., Taylor K., Whitehead S., Barrell B.G.  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z95436; CAB08826.1; -.  
DR EMBL; AE007171; AAK48084.1; -.  
DR TIGR; MT3723; -.  
DR TubercuList; Rv3621c; -.  
DR InterPro; IPR000030; Microbac\_PPE.  
DR Pfam; PF00823; PPE; 1.  
DR Hypothetical protein; Complete proteome.  
SQ SEQUENCE 413 AA; 40679 MW; DAFFE8917A900A3 CRC64;  
  
Query Match 36.0%; Score 702; DB 16; Length 413;  
Best Local Similarity 39.1%; Pred. No. 1.5e-31;  
Matches 168; Conservative 61; Mismatches 145; Indels 56; Gaps 7;  
  
QY 1 MVDGALPPEINARBYAGGSGASLVAAQOMDVSADLFSASAFQSVVWGLTVGSWIG 60  
DB 1 MVDGALPPEINARBYAGGSGGPMPLAAAEALAEELTGTASTYDALITGLADGPWGQ 60  
QY 61 SSAGLWVAASPYVAMSVTGAQELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 120  
DB 61 SSASAVVAATPVAVNLSTAGAGAEQAGSQAVALAAAYEAFAATVPPELIANRALVLA 120  
QY 121 LATNLGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLLPFEAEPMTSAGG 180  
DB 121 LATNLGONTAAIAATEAYAEWMAODAAAFGYAAGASAAAT-QLSFENPAQTINPAG 179  
QY 121 LLEQAAVEEASSTAAANQUMNVPAQLQOLAQPTGTTSSSLGSLMKTVSHPSPISN 240  
DB 181 LLEQAAVEEASSTAAANQUMNVPAQLQOLAQPTGTTSSSLGSLMKTVSHPSPISN 240  
QY 180 LMSQAASVQOAVSAGANNAQALTDIPAL-----FGLSGFTNEPMTLDLKG 226  
DB 241 MYSMANNHSMNTSVSMNTTSSMLKGFAPAA---AAQAVOT-----280  
QY 227 AGLTHTHTSSDSSGLIVGVLGDFVQVYGSALDASVAMDFGKQVSPARLMVTOFQD 286  
DB 281 -----AAQGVAMSSLSGSSGLGGVAAANTGRAASVGSLSVPOA 323  
QY 287 YFGLAHLDFKWASEGAKAAGEAKALPAVPAIPASGL-SGVAGAVGQAASVGLKVPAY 345  
DB 324 WAAANQAVTPPARALTLTSLTAERGPGQMLGGLVQGWGARAAGGLSVL--RVPRP 381  
QY 346 WHTTTTAAASDAVTAASNGAALAAAGSTHAFGMPL--NGSGAGAFNNFAAPRYGPR 403  
DB 382 YVMPHSPAG 391  
QY 382 YVMPHSPAG 391

DB 404 TVIAQPPAG 413  
  
RESULT 12  
QY 08VJWO PRELIMINARY; PRT; 694 AA.  
AC 08VJWO;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE PPE family protein.  
GN MT1857.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriinae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxId=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE007044; AAK46130.1; -.  
DR TIGR; MT1857; -.  
DR InterPro; IPR000568; ATPase\_Aeub.  
DR InterPro; IPR000030; Microbac\_PPE.  
DR Pfam; PF00823; PPE; 1.  
DR PROSITE; PS00449; ATPase\_A; 1.  
SQ SEQUENCE 694 AA; 73206 MW; 2FDDAD94290A07359 CRC64;  
  
Query Match 35.1%; Score 683.5; DB 16; Length 694;  
Best Local Similarity 41.2%; Pred. No. 3.1e-30;  
Matches 166; Conservative 58; Mismatches 122; Indels 57; Gaps 10;  
  
QY 2 VDFGALPPEINARBYAGGSGASLVAAQOMDVSADLFSASAFQSVVWGLTVGSWIG 61  
DB 227 VDFGALPPEITSGEWLIGGAGPMPLAAAVAGDLAELOSMASVSYVEGMASESMLG 286  
QY 62 SAGLWVAASPYVAMSVTGAQELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 121  
DB 287 SSAGLWVAASPYVAMSVTGAQELTAQVRVAAAAYETAYGLTVPPVIAENRAELMT 121  
QY 122 IATNLGONTPAIAVNEAEYGEWMAODAAAFGYAATATATATLLPFEAEPMTSAGG 181  
DB 347 VATNIFGONTAAIAATEAYAEWMAODTMAAMGYASSSATA-SRLTFPTAPPTNPISL 405  
QY 182 LLEQAAVEEAS-----DTAANQUMNVPAQLQOLA---OPTQ-----216  
DB 406 ACPAAATGATLALAGCTNAVTALSSAAQPFIDITLLGGLATLSTGYTQUMGLINA 465  
QY 217 -GTPSSSLGSLMKTVSHPSPISNMYSMANNHSMNTSVSMNTTSSMLKGFAPAA 274  
DB 466 IPEPGATTYQNLFTVA-----ANVTKFTWADNAPSAPMLGTEFVF-----WOPPPA 515  
QY 275 AAVOTPAQNGRAMSSLSG-----SSGLGGVAAANTGRAASVGSLSVPOAANAQ 329  
DB 516 PEIPK-----SSLAGAGLRSGLSAGLAHAASALGQANLVGDLSSVPSWASATP 565  
QY 330 AVTPAARALPLTSLTA-AERGPGQMLGGLVQGWGARAAGGL 371  
DB 566 AVRLVANTPLPATSLAAAPATQIPANLLGOMALGSM---TGGAL 605  
  
RESULT 13  
ID 053958 PRELIMINARY; PRT; 468 AA.  
AC 053958;  
DT 01-JUN-1998 (Tremblrel. 06, Created)



DT 01-JUN-1998 (TReMBLrel. 06, last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)  
 DE PPE-family protein.  
 GN RV1809 OR MTW049.31.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 RX STRAIN=H37Rv;  
 RC MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 DR EMBL; AL022021; CAI1730.1; -  
 DR Tuberculist; Rv1809; -  
 DR InterPro; IPR000568; ATPsynb\_Asub.  
 DR InterPro; IPR00030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 DR PROSITE; PS00449; ATPase\_A; 1.  
 DR Complete proteome.  
 KW KW  
 SQ SEQUENCE 468 AA; 47156 MW; 14D05073BF68028A CRC64;  
 Query Match 35.0%; Score 682.5; DB 16; Length 468;  
 Best Local Similarity 40.9%; Pred. No. 2,2e-30;  
 Matches 164; Conservative 59; Mismatches 125; Indels 53; Gaps 9;  
 QY 2 VDFGALPPEINSAARMYAGPSASLVAAQWDSVADLFSASAFOGVWGLTVGSWIGS 61  
 DB 1 MDRFLQPEITSGEMVIGPGAGPMLAAAVMDGIAELQSWASIVSGMASESWLGP 60  
 QY 62 SAGLWVAASPYVAMSVTAGQALTAQVRVAAAATETAYGLTVPPVIAENRAELMITL 121  
 DB 61 SSAGMAAAAPYVWMSGTSAQAADQAAAVAAETFAAVVPPPOIAANNSQILSL 120  
 QY 122 IATNLGONTPAIVNBAEYGEWMAODAAAFGYAAATATATLTFEEBAPENTSAGL 181  
 DB 121 VANNIFGONTAAIAATEAEGEMMAQDTMAVFGYASSATP-SRLTFTTAPPOTTNSGL 179  
 QY 182 LEQAAAVEAS-----DTAANQLMNNVPOALQOLAQPTQGTTPSSKLGGL- 227  
 DB 180 AGQAAAGQADALASGTNAVTTLASAAQPPDIITLLQGLA--TLSTQYQLMGQLI 237  
 QY 228 -----WKTVSHRSPISNMVSMANNHNSMTNSGVSMNTLSSMLKGFAPAAAQ 276  
 DB 238 NAIFGPTGATTYQVVFYTAANVTKEFTWANDAMSAPMLGMEFKVF-----MQPPAPE 291  
 QY 277 AVQTAAGNVRANSSIGSSIG-----SSGIGGVAAVNLGRAASVGSISVPOAAAQAV 331  
 DB 292 IPR-----SSLDAGLGLASGLSALHAAASAGIGQANLVLDGSLVSPSWASATPAV 341  
 QY 332 TPAARALPLTSLTSA-AERGFQMLGGLPYQMGARAGCGT 371  
 DB 342 RLVAANTLIPATSLAAAPATQIPANLIGMALGSM---TGAL 379  
 RESULT 14  
 ID P95190 PRELIMINARY; PRT; 380 AA.  
 AC P95190;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Hypoetical 38.0 kDa protein (PPE family protein).

GN RV3136 OR MTQ03A2.22C OR MTJ3221.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 RX STRAIN=H37Rv;  
 RC MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE007137; AAK47561.1; -  
 DR TIGR; MT3221; -  
 DR Tuberculist; RV3136; -  
 DR InterPro; IPR00030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 DR Hypoetical protein; Complete proteome.  
 KW KW  
 SQ SEQUENCE 380 AA; 37979 MW; 345B1EF9C9A84AA CRC64;  
 Query Match 34.6%; Score 675; DB 16; Length 380;  
 Best Local Similarity 43.1%; Pred. No. 4,4e-30;  
 Matches 172; Conservative 56; Mismatches 143; Indels 28; Gaps 12;  
 QY 2 VDFGALPPEINSAARMYAGPSASLVAAQWDSVADLFSASAFOGVWGLTVGSWIGS 61  
 DB 1 MDRFLQPEITSGEMVIGPGAGPMLAAAVMDGIAELQSWASIVSGMASESWLGP 60  
 QY 62 SAGLWVAASPYVAMSVTAGQALTAQVRVAAAATETAYGLTVPPVIAENRAELMITL 121  
 DB 61 AASMSAVTAPYICWLYTTAEKTOQTALQARAAALFEQAVYAMTLPPIVAAARLQIAL 120  
 QY 122 IATNLGONTPAIVNBAEYGEWMAODAAAFGYAAATATATLTFEEBAPENTSAGL 181  
 DB 121 IATNFGONTAAIAATEAQAEMWMAODAAAMGYATAA-AAALITLPPFSPROTTPAGL 179  
 QY 182 LEQAAAVEAS-----DTAANQLMNNVPOALQOLAQPTQGTTPSSKLGKTKTSPHSP 238  
 DB 180 TAAAPASQATDPLSL--LLETVOALQALTTPSFIPEDFTLDAIPAGYATGVTQDVE 237  
 QY 239 SNNVSM--ANNHNSMTNSGVSMNTLSSMLKGFAPAAAQAVQTAONGVRANSSIGSSL 296  
 DB 238 SFVAGTIGAESNIGLNVGDENPAEVTPODGFGLGELVSA---TSPGGGVSA--SGAG--- 289  
 QY 297 GSSGIGGVAAVNLGRAASVGSISVPOAAAQAVTPAARALPLTSL--TSAERG-PCQ 353  
 DB 290 GAASVGNVTLASVGRANSGISVSPSWAAPSRTVSALSACLTTLPGTDAVEHGMPG- 348  
 QY 354 MLCGLPVGQMGARAGGLSGVL-RVPPRPVWHPSPAAG 391  
 DB 349 -VPGVPV-----AAGRASGVLPRIYGRVLTVAHHPAAG 380



```

RESULT 15
033310 PRELIMINARY; PRT; 462 AA.
AC 033310;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DE 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE PPE-family protein.
GN RV2768C OR MT2838 OR MTV002.33C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
(2)
RN RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleisichmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwin M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Knouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL008967; CAA15564.1; ALT_INIT.
DR TIGR; MT838; -.
DR Tuberculist; RV2768C; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ
SEQUENCE 462 AA; 46307 MW; A65F35B59A9C8773 CRC64;

Query Match 34.3%; Score 668.5; DB 16; Length 462;
Best Local Similarity 40.8%; Pred. No. 1.3e-29;
Matches 161; Conservative 61; Mismatches 156; Indels 17; Gaps 7,

QY 2 VDFGALPPEINSAARMYAGPGSASLVAAQWDSVASDLPFAASAFOSVVGTLVGSWIGS 61
DB 69 MDFGALPPEINSTRMYAGAGAPLMAAGATWGLAVELSTTASSVSVIMQLTTEQWLGP 128
QY 62 SAGLMTAAASPYVAMSVTAQAEITAAQYVAAAAYETAYGTLTUPPVIATENRAELMTL 121
DB 129 ASMSMVAAQPYLAULTYAESAAHAAQAQMASAAAFAMTVPFAEVAANRALIAAL 188
QY 122 IATNLLGONTPAIVAEAEYGEWMAODAAAFGYAATATATATLTPFEAPEMTSAGGL 181
DB 189 VATNVGQNTPAIMATEAHYGEWMAODALAMGYAASSA-AAGRLNPLTTPSOTANMAGL 247
QY 182 LEQAAAVEEASDPTAANQ-----LMMNVQALQQLAOPTGTTSSSKGLMKTWSP--H 234
DB 248 ACGAAAVSHAAASTVQVGLSLISNLPVAVMGFASPLTSADAAGLGIIQDIEELLG 307
QY 235 RSPDISMVMANNHMTSGVSMNTTLLSMLKGFAPAAAQAVQTPAONGV---RAMSS 291
DB 308 ITFVQNAINGAVN---TTAMFWMATIPNAVPLGHAFAALNPATVTAADAVPAAAAAAG 363
QY 292 LGSLSGSSGLGG-GVAANLGRAASVGLSLVPGQMAAANAQAVTPAARALPLTSLTSAERG 350

```

```

DB 364 LAHTVTPVGVGASLTLASLGEASVVGSLVPAGWSTPAADAMTSGTTALEGSGWAVPEEAG 423
QY 351 PQMTGLGFPVGMGARAAGGLSGVLRAVPPRPVMP 385
DB 424 PYAAMPGM-AGISGAAGAGAYAGPRYGFPIVMP 457

```

Search completed: November 21, 2003, 17:05:32  
 Job time : 43 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:02:14 ; Search time 20 Seconds

(without alignments)  
1880.098 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDFGALPPEINSARMYAGP.....SGVLKVPPEPYVMPHSPAAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR\_76:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Query Match	Length	ID	Description
1	1949	100.0	391	2 B70608	probable PPE prote
2	1656.5	85.0	396	2 H70741	probable PPE prote
3	1583	81.2	393	2 C70568	probable PPE prote
4	779.5	40.0	393	2 G70929	probable PPE prote
5	751	38.5	409	2 A70932	probable PPE prote
6	737.5	37.8	403	2 H70931	probable PPE prote
7	736	37.8	423	2 B70931	probable PPE prote
8	731	37.5	421	2 H87056	PPE-family protein
9	705	36.2	408	2 G70925	probable PPE prote
10	702	36.0	391	2 B70625	probable PPE prote
11	702	36.0	413	2 F70560	probable PPE prote
12	688.5	35.3	463	2 C70931	probable PPE prote
13	682.5	35.0	468	2 B70932	probable PPE prote
14	675	34.6	380	2 A70646	probable PPE prote
15	668.5	34.3	394	2 G70881	probable PPE prote
16	667.5	34.2	385	2 H70503	probable PPE prote
17	635	32.6	350	2 H70929	probable PPE prote
18	633.5	32.5	365	2 E70929	probable PPE prote
19	615.5	31.6	402	2 A70882	probable PPE prote
20	603	30.9	423	2 C70582	probable PPE prote
21	597.5	30.7	391	2 D70922	probable PPE prote
22	589	30.2	406	2 E70675	probable PPE prote
23	588	30.2	394	2 A70504	probable PPE prote
24	575.5	29.5	391	2 A70663	probable PPE prote
25	495.5	25.4	3300	2 D70575	probable PPE prote
26	473	24.3	180	2 G70834	probable PPE prote
27	452.5	23.2	3716	2 E70969	probable PPE prote
28	448.5	23.0	580	2 G70570	probable PPE prote
29	444	22.8	346	2 H70874	probable PPE prote

30	443	22.7	963	2 B70524	probable PPE prote
31	441	22.6	487	2 C70830	probable PPE prote
32	439.5	22.6	655	2 A70931	probable PPE prote
33	429.5	22.0	479	2 D70676	probable PPE prote
34	428.5	22.0	1053	2 B70987	probable PPE prote
35	426.5	21.9	443	2 C70780	probable PPE prote
36	423.5	21.7	2523	2 F70846	probable PPE prote
37	422.5	21.7	615	2 E70663	probable PPE prote
38	418	21.4	678	2 A70762	probable PPE prote
39	418	21.4	3157	2 B70969	probable PPE prote
40	415.5	21.3	618	2 H70552	probable PPE prote
41	415	21.3	582	2 F70675	probable PPE prote
42	414.5	21.3	987	2 E70808	probable PPE prote
43	412	21.1	590	2 E70946	probable PPE prote
44	412	21.1	645	2 F70825	probable PPE prote
45	409.5	21.0	409	2 A70647	probable PPE prote

#### ALIGNMENTS

##### RESULT 1

B70608 probable PPE protein - Mycobacterium tuberculosis (strain H37Rv)

Db 361 GQMGARAGGGLSVLRVPPRYVMPHSPAAG 391

## RESULT 2

H70741 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: H70741

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

R/Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: H70741

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-396 <COL>

A/Cross-references: GB:Z75555; GB:AL123456; NID:g3261608; PIDN:CAA99966.1; PID:e250360;

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: PPE

Query Match 85.0%; Score 1656.5; DB 2; Length 396;

Best Local Similarity 85.1%; Pred. No. 6.4e-91;

Matches 338; Conservative 18; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWMDVSADLFSASAFQSVVWGLTVGSMIG 60

Db 1 MVDGALPPEINSARMYAGPGSASLVAAQWMDVSADLFSASAFQSVVWGLTVGSMIG 60

QY 61 SSAGLMAAASPVAAMSVTAAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMT 120

Db 61 SSAGLMAAASPVAAMSVTAAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMT 120

QY 121 LIATNLGONTPTAIVAEVGEWMAODAAAFGYAATATATATLTPFEAPEMTSAGG 180

Db 121 LIATNLGONTPTAIVAEVGEWMAODAAAFGYAATATATATLTPFEAPEMTSAGG 180

QY 181 LLEQAAVEBASPAAANQLMNNVPOALQOAPGTPSSKLGIMKTVPSPRSPISN 240

Db 181 LLEQAAVEBASPAAANQLMNNVPOALQOAPGTPSSKLGIMKTVPSPRSPISN 240

QY 241 MVSANNNHSMSTNGSVMTNTLSMLKGFAPAAAQAVETAAQNGVAMSSLSGSL 296

Db 241 MVSANNNHSMSTNGSVMTNTLSMLKGFAPAAAQAVETAAQNGVAMSSLSGSL 296

QY 297 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 356

Db 297 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 356

QY 300 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 359

Db 300 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 359

QY 357 GLPVQMGARAG--GGLSGVLRVPPRYVMPHSPAAG 391

Db 357 GLPVQMGARAG--GGLSGVLRVPPRYVMPHSPAAG 391

QY 360 GLPLGLTNSGCGGCGVSNALRMPRPAVMPRVPDPAAG 396

Db 360 GLPLGLTNSGCGGCGVSNALRMPRPAVMPRVPDPAAG 396

## RESULT 3

C70568 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: C70568

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

R/Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: C70568

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-393 <COL>

A/Cross-references: GB:Z95390; GB:AL123456; NID:g3261766; PIDN:CAB08702.1; PID:e316074; I

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: PPE

Query Match 81.2%; Score 1583; DB 2; Length 393;

Best Local Similarity 81.5%; Pred. No. 1.4e-86;

Matches 322; Conservative 20; Mismatches 47; Indels 6; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWMDVSADLFSASAFQSVVWGLTVGSMIG 60

Db 1 MVDGALPPEINSARMYAGPGSASLVAAQWMDVSADLFSASAFQSVVWGLTVGSMIG 60

QY 61 SSAGLMAAASPVAAMSVTAAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMT 120

Db 61 SSAGLMAAASPVAAMSVTAAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMT 120

QY 121 LIATNLGONTPTAIVAEVGEWMAODAAAFGYAATATATATLTPFEAPEMTSAGG 180

Db 121 LIATNLGONTPTAIVAEVGEWMAODAAAFGYAATATATATLTPFEAPEMTSAGG 180

QY 181 LLEQAAVEBASPAAANQLMNNVPOALQOAPGTPSSKLGIMKTVPSPRSPISN 240

Db 181 LLEQAAVEBASPAAANQLMNNVPOALQOAPGTPSSKLGIMKTVPSPRSPISN 240

QY 241 MVSANNNHSMSTNGSVMTNTLSMLKGFAPAAAQAVETAAQNGVAMSSLSGSL 296

Db 241 MVSANNNHSMSTNGSVMTNTLSMLKGFAPAAAQAVETAAQNGVAMSSLSGSL 296

QY 297 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 356

Db 297 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 356

QY 300 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 359

Db 300 GSSGLGGVAAANIGRAASVGSLSVPQMAAANQAVTPPAARALPLTSLTSAERGPQMIG 359

QY 357 GLPVQMGARAGGGLSGVLRVPPRYVMPHSPAAG 391

Db 357 GLPVQMGARAGGGLSGVLRVPPRYVMPHSPAAG 391

QY 360 GLPLGLTNSGCGGCGVSNALRMPRPAVMPRVPDPAAG 393

Db 360 GLPLGLTNSGCGGCGVSNALRMPRPAVMPRVPDPAAG 393

## RESULT 4

G70929 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: G70929

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

R/Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: G70929

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-393 <COL>

A/Cross-references: GB:AL022021; GB:AL123456; NID:g3250699; PIDN:CAA17711.1; PID:e125460

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: PPE

Query Match 40.0%; Score 779.5; DB 2; Length 393;

Best Local Similarity 43.7%; Pred. No. 4.3e-39;

Matches 179; Conservative 65; Mismatches 129; Indels 37; Gaps 10;

QY 2 VDEGALPPEINSARMYAGPGSASLVAAQWMDVSADLFSASAFQSVVWGLTVGSMIG 61

Db 1 MVDGALPPEINSARMYAGPGSASLVAAQWMDVSADLFSASAFQSVVWGLTVGSMIG 60

QY 62 SAGLMAAASPVAAMSVTAAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMT 121

Db 61 SAGLMAAASPVAAMSVTAAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMT 120

QY 122 IATNLLGONTPAIYVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAEEMTSAGL 181  
 Db 121 ISTNVFGQNTSAIAAEAYGEWMAQDSAAWYAVASSASASA-VTFSTPPQIANPTAQ 179  
 QY 182 LEQAAVAEASDTAA--NQLMNNVPAOLQOLAOPFGTTPSSKLGKLMK----- 229  
 Db 180 GTQAAAVATVTAQSTLTLEMTGTLPNALQSLTSPLOS--NGPLSLWLIQIIFGTPTNPT 238  
 QY 230 -----TVSPRSPISNNVSMANNMMSMTNSGVSMTNTLSMLKGFAPAAAQAQVTA 282  
 Db 239 SISALLTDLQPAVSPFYNTGEGLPYFISGNNFISQAKTL-GIGSAAPAVA-----NA 292  
 QY 283 QNGVAMSSLSGSSGCGVVAANTGRAASVGSISVPOAANA-ANQAVTPPAARALPLT 341  
 Db 293 GAAAGLPELGGMLG---GCPVAAAGLGNAAASVKSLSVPPVMSGRLPGSVTGAAPLPS 348  
 QY 342 SITSAAERPGOMLGLPVGOMGARAGGLSGVLRVPPPPVPHSPAAG 391  
 Db 349 TVSAAPPEARPGSLGLPL---AGAGGAGAGP-RYGFPRPTWAPRPFAG 393

## RESULT 5

A70932  
 Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: A70932  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: A70932  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-409 <COL>  
 A:Cross-references: GB:AL022021; GB:AL123456; NID:g3250699; PIDN:CA11772.1; PID:e125461  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: PPE

Query Match 38.5%; Score 751; DB 2; Length 409;  
 Best Local Similarity 44.7%; Pred. No. 2.2e-37;  
 Matches 177; Conservative 53; Mismatches 138; Indels 28; Gaps 10;  
 QY 2 VDFGALPPEINSARMYAGPGSASLVAAQOMDVSADLPSAASAFQSVVWGLTVGSWIGS 61  
 Db 1 MFGALPPEINSGRMVAGPGSGPLAAALAAALAEALYSAASVSGSTTEGLTVAPMKGP 60  
 QY 62 SAGLVAAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPPIYANRAELMTL 121  
 Db 61 SSITVAAAVAPVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPPIYANRAELMTL 120  
 QY 122 IATNLLGONTPAIYVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAEEMTSAGL 181  
 Db 121 VATNIFGONTPAIATEAHYAEWMAQDAAMFGYAAATATATATLLPFEAEEMTSAGL 179  
 QY 182 LEQAAVAEASDTAA-----NQLMNNVPAOLQOLAOPFGTTPSSKLGKLMKTVSPH 234  
 Db 180 AAQSAVVAQAAGAAASDITAOISILPSTLOSIA--TTATATSAAG--WDTV--- 232  
 QY 235 RSPISNNVSMANNMMSMTNSGVS---MTNTLSMLKGFAPAAAQAQVTAQNGVRAM 289  
 Db 233 IQSITITLANTLPYSIILGAIPEGWMTFCQIIGLAQNAAGVALLCPKKAAGALSP 292  
 QY 290 SSL-SSISGS-SGLGGVVAANTGRAASVGSISVPOAANAANQAVTPPAARALPLTSTSA- 346  
 Db 293 APLRGVYIGDITPLTGGATGTATRAIYVGSLSVPGWMAAIVMRAVASVLGTGAAPL 352  
 QY 347 AERPGOMLGLPVGOMGARAGGL-----SGVLRV 377

Db 353 AAEPGALFGEWALSLAGRALAGTAVRSGANARV 388

## RESULT 6

H70931  
 Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: H70931  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: H70931  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-403 <COL>  
 A:Cross-references: GB:AL022021; GB:AL123456; NID:g3250699; PIDN:CA11772.1; PID:e125461  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: PPE

Query Match 37.8%; Score 737.5; DB 2; Length 403;  
 Best Local Similarity 42.6%; Pred. No. 1.3e-36;  
 Matches 181; Conservative 60; Mismatches 123; Indels 61; Gaps 12;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQOMDVSADLPSAASAFQSVVWGLTVGSWIGS 61  
 Db 5 LDFATLPEINSARMYAGPGSASLVAAQOMDVSADLPSAASAFQSVVWGLTVGSWIGS 64  
 QY 62 SAGLVAAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPPIYANRAELMTL 121  
 Db 65 ASASMTAAAPVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPPIYANRAELMTL 124  
 QY 122 IATNLLGONTPAIYVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAEEMTSAGL 181  
 Db 125 IATNVLGONTPAIATEAHYAEWMAQDAAMFGYAAATATATATLLPFEAEEMTSAGL 183  
 QY 182 LEQAAVAEASDTAA-----NQLMNNVPAOLQOLAOPFGTTPSSKLGKLMKTVSPH 234  
 Db 184 AAQSAVVAQAAGAAASDITAOISILPSTLOSIA--TTATATSAAG--WDTV--- 232  
 QY 235 RSPISNNVSMANNMMSMTNSGVS---MTNTLSMLKGFAPAAAQAQVTAQNGVRAM 289  
 Db 234 WLDKRLMALDPR-----SNFMNTIASGFLPFSNTIAPFLGLGVAADAAGDV 293  
 QY 279 QTAQNGVRAMSSLSGSSL-----GSSGIGGVVAANTGRAASVGSISVPOAANAANQAVTP 333  
 Db 294 LGEATSG-----GLGGALVAPLGSAGLGTVAAGLGNAAATGTLSVPPSWTAAPPLASP 348  
 QY 334 AARAL---PLTSLTAAERPGOMLGLPVGOMGARAGGLSGVLRVPP---RYVMMH 386  
 Db 349 LGSALGCTPMAAPPAVAAAG---MPGMPFGTMSGGGFG---RAVQYGFRRNFVAR 398  
 QY 387 SPAAG 391  
 Db 399 PPAAG 403

## RESULT 7

B70931  
 Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: B70931  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998

Query Match 37.5%; Score 731; DB 2; Length 421;  
Best Local Similarity 41.1%; Pred. No. 3.4e-36;  
Matches 174; Conservative 55; Mismatches 160; Indels 34; Gaps 8;

[illegible]

Db 180 AGGAAATVSTVPLATTAAPVLLQOLSTSTSLIPWVSALQOVLAEHLGLTFDNRMTIV 239  
QY 225 -----GGLMKTVSPHRSPISSNMVSNMNMNTNGSVSTNTLSSMLKCFAPAAAA 275  
Db 240 RLIGISYFDEGL-----LQFASLTAQOAIPEGPGAG--DSGSSVLSWGWGPTTIRA 287  
QY 276 QAVQTAAQNGVRAMSSL--GSSLSGS-----SGLGGGVAAVNGRAASVGSIS 319  
Db 288 -----GPRASPSVAGGAGVGVQOTQPPYMYMALDRESIGGSVSAALGKSSAGSIS 338  
QY 320 VPDMAAANAQAVTPPARALP---LTSLSAERGPQMTGLGVQGVGAGAGGSLGVL 376  
Db 339 VPPDMAARARWAMPARLPDGDVDTALRGTAENA---LIRGPFMAAGSGTGGGF--VHK 393  
QY 377 VPPRPVYMHSPAG 391  
Db 394 YGFRLLAVMORPPAG 408

## RESULT 10

B70625  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: B70625  
R:Coile, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; PMID:9825987; PMID:9634230  
A:Accession: B70625  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-391 <COL>  
A:Cross-references: GB:292539; GB:AL123456; NID:g3261714; PIDN:CAB06873.1; PID:e304546.  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: PPE

Query Match 36.0%; Score 702; DB 2; Length 391;  
Best Local Similarity 42.3%; Pred. No. 1.6e-34;  
Matches 166; Conservative 62; Mismatches 150; Indels 14; Gaps 7;  
QY 2 VDFGALPPEINSARMYAGPGSASLVAAQMWDSVASDLPSAASAFOSVVMGLTVGSWIGS 61  
Db 1 MDFGALPPEINSARMYAGPGSASLVAAQMWDSVASDLPSAASAFOSVVMGLTVGSWIGS 60  
QY 62 SAGLWMAASPYVAMSVTAQAEELTAQVRYAAAAETAYGLTVPPIAENRAELMTL 121  
Db 61 ASMAWMAAOPYLAWMTLYAEAAAHASQAMASAAAYEAAYMTVPEVVAANRALLAAL 120  
QY 122 IATNLLGONTPAIVNEAEYGEWMAODPAAMFCYATATATATLLEPEAEPMTSAGL 181  
Db 121 VATTNVLGINTPAIMATEALYAEKMAODALAMYGYAASG--AAGMLQPLSPSQTTPGGL 179  
QY 182 LEQAAAEEASDPTAAANQ-----LMNNVPOALQOLAOPTQGTTPSSKLGGLMKTVSPHRS 236  
Db 180 AAGSAAVGSAAATTAANQSVADLISLIPNAVSGLASPTSVLSDTGLSGLIADIDALLA 239  
QY 237 P---ISNMVSMANNHMS--MTNSGVSMNTLSSMLKGFAPAAAAQAVTAAQNGVRAMSSIG 293  
Db 240 TPFVANNITNSAVNTAAVYNAAIPTAIFLANLNSGAPVAILGALTEAEG--AASAAA 296  
QY 294 SSLSSGSGGGAANNGRAASVGSLSVPOMAAANAQAVTPPARALPLTSLTSAERGPQ 353  
Db 297 AGLAASVTPAGLGSISGEKTLVGRSLVPAAMKSTAAATTAGATLARGSGVTVAABEA--GP 355  
QY 354 MLGGLPVGVMGARAGGSLGVLVPPRPVYMP 385  
Db 356 VTGMMP--GMAAAGTGGAAYAGPRYGFKPTVMP 386

## RESULT 11

F70560  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: F70560  
R:Coile, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; PMID:9825987; PMID:9634230  
A:Accession: F70560  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-413 <COL>  
A:Cross-references: GB:295436; GB:AL123456; NID:g3261770; PIDN:CAB08826.1; PID:e31775.1  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: PPE

Query Match 36.0%; Score 702; DB 2; Length 413;  
Best Local Similarity 39.1%; Pred. No. 1.7e-34;  
Matches 168; Conservative 61; Mismatches 145; Indels 56; Gaps 7;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQMWDSVASDLPSAASAFOSVVMGLTVGSWIG 60  
Db 1 MVDGALPPEINSARMYAGPGSASLVAAQMWDSVASDLPSAASAFOSVVMGLTVGSWIG 60  
QY 61 SAGLWMAASPYVAMSVTAQAEELTAQVRYAAAAETAYGLTVPPIAENRAELMT 120  
Db 61 SGAAMVMAATQVAVMLRSTAGAQAGSQAQAAVAAASVYEAAPFATVPPEIAANRALMA 120  
QY 121 LATNLLGONTPAIVNEAEYGEWMAODPAAMFCYATATATLLEPEAEPMTSAGL 180  
Db 121 LATNLLGONTPAIVNEAEYGEWMAODPAAMFCYATATATLLEPEAEPMTSAGL 179  
QY 181 LEQAAAEEASDPTAAANQ-----LMNNVPOALQOLAOPTQGTTPSSKLGGLMKTVSPHRSPISN 240  
Db 180 LASQAASVQAAGVGAANQALTDIPKAL-----FGLSIFINPEPMLTDLCK 226  
QY 241 MYSMANNHMSMTNGSVSMNTLSSMLKGFAPAA---AAQAVT----- 280  
Db 227 ALGLTGHWTSSDGSGLIYGVGLGDEYQGVTSABLDASVAMDTFGKWSPARLMTQFKD 286  
QY 281 -----AAQNGVRAMSSLSGSSLSGSGLGGVAAVNLGRAASVGSLSVPOA 323  
Db 287 YFGLAHDLPKMASGAKAAGEAKALPAAVPAIPFAGL--SGVAGAVGQAASVGLKVPAY 345  
QY 324 WAAANAQAVTPPARALPLTSLTSAERGPQMTGLGVQGVGAGAGGSLGVL--RVPP 381  
Db 346 WATTPPASPAVLAASNGGAAAAAEGSTHACGNPL--MSGAGRAFPNFAAPRGFRP 403  
QY 382 YVMPHSPAG 391  
Db 404 TVIAQPPAG 413

## RESULT 12

C70931  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: C70931  
R:Coile, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230  
 A/Accession: C70931  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Strains: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-463 <COL>  
 A/Cross-references: GB:AL020201; GB:AL123456; NID:93250699; PIDN:CAA17723.1; PID:e125461  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Genes: PPE

Query Match 35.3%; Score 688.5; DB 2; Length 463;  
 Best Local Similarity 42.8%; Pred. No. 1.2e-33;  
 Matches 174; Conservative 48; Mismatches 136; Indels 49; Gaps 11;

QY 2 VDEGALPEINSAARMYAGPGSASLVAAAOQWDSVASDLFSAASAFQSVWGLTVGSWIGS 61  
 Db 1 MDGVLPEINSGMYAGPGSAPMLAAAMDGATLQSTAAAYGVISVLT-GVWSGQ 59  
 QY 62 SAGLVAAASPYVAMSVTAGQAEITAAQVRVAAAAYETAVGLTPPPVIAENRAELMIL 121  
 Db 60 SSGTMAAAAAPYVAMSVTAGQAEITAAQVRVAAAAYETAVGLTPPPVIAENRAELMIL 119  
 QY 122 IATNLLGONTPAIVNAEYGEWMAODAAAFGAAATATATATLTFEEAPEMTSAGL 181  
 Db 120 AATNIFGONTGALIAAARAEWMAODAAAFGAAATATATATLTFEEAPEMTSAGL 178  
 QY 182 LEQAAVEASDTPAAANQMLNNVPOALQLOAOTPGTTPSSKLGIMKTYS--PHRSP- 238  
 Db 179 ATQGVAAQVAGASAGN-ASLSVEUEFLA--TAGTNYKTVASLMMNATVGPVYASSVY 235  
 QY 239 -----SNVSMANNHSMNTSGVSMNTTSLSMKGPAPAAAQVQTAQNGVRA 288  
 Db 236 NSMLGLGFAASKVLPANDTVITSTIFGVOQFQKFPNFTVPNDLIRK----- 283  
 QY 289 MSSLGSSLG-----SSGLGCG--GVANLGRASVGSLSVPQMAAAAOAVTPAARLPL 340  
 Db 284 -SALGAGLGRSAISSGLGSTAPASISGASQAGSGVSPPSVAAATPAIRVAAVSS 342  
 QY 341 TSLTS--AERGPQOML-----GGLPYQMGARAGGGLGSLTVR 377  
 Db 343 TGLQAVPAAAISSGLSLSQMAASVAGALGGAARATGFLGGGRV 389

# RESULT 13

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: B70932

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: B70932

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-468 <COL>

A/Cross-references: GB:AL020201; GB:AL123456; NID:93250699; PIDN:CAA17730.1; PID:e125462

A/Experimental source: strain H37RV

C/Genetics:

A/Genes: PPE

Query Match 35.0%; Score 682.5; DB 2; Length 468;  
 Best Local Similarity 40.9%; Pred. No. 2.8e-33;  
 Matches 164; Conservative 59; Mismatches 125; Indels 53; Gaps 9;

QY 2 VDEGALPEINSAARMYAGPGSASLVAAAOQWDSVASDLFSAASAFQSVWGLTVGSWIGS 61  
 Db 1 MDGVLPEINSGMYAGPGSAPMLAAAMDGATLQSTAAAYGVISVLT-GVWSGQ 59

QY 62 SAGLVAAASPYVAMSVTAGQAEITAAQVRVAAAAYETAVGLTPPPVIAENRAELMIL 121  
 Db 61 SSGTMAAAAAPYVAMSVTAGQAEITAAQVRVAAAAYETAVGLTPPPVIAENRAELMIL 120  
 QY 122 IATNLLGONTPAIVNAEYGEWMAODAAAFGAAATATATATLTFEEAPEMTSAGL 181  
 Db 121 AATNIFGONTGALIAAARAEWMAODAAAFGAAATATATATLTFEEAPEMTSAGL 179  
 QY 182 LEQAAVEASDTPAAANQMLNNVPOALQLOAOTPGTTPSSKLGIMKTYS--PHRSP- 227  
 Db 180 AGQAAATGQATLALASGTNAVTTALSSAAQFPDITLTLQGLA--TLSTQYTLQMGQL 227  
 QY 228 -----MKTSPHRSPISSNMVSMANNHSMNTSGVSMNTTSLSMKGPAPAAAQ 276  
 Db 238 NALFPGTGAATTVQNVFVTAANVTKFSTWANDAMSANLGMTKRVF-----WQPPAPE 291  
 QY 277 AVQTAQNGVRAMSSISGSSLG-----SSGLGCGVAAANLGRASVGSLSVPQMAAAAOAV 331  
 Db 292 IPR-----SSLGAGLGRSAGLSAGLAHAASAGLQANLVGDLSPPSWASATPAV 341  
 QY 332 TPAAALPLTSLTSA-AERGPQOMLGLPYQMGARAGGGL 371  
 Db 342 RLVAATLPLTSLTSAAPATQIPANLQOMLGSW--TGAL 379

# RESULT 14

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: A70646

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: A70646

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-380 <COL>

A/Cross-references: GB:Z83867; GB:AL123456; NID:93261695; PIDN:CAB06278.1; PID:e291015; I

A/Experimental source: strain H37RV

C/Genetics:

A/Genes: PPE

Query Match 34.6%; Score 675; DB 2; Length 380;  
 Best Local Similarity 43.1%; Pred. No. 6.1e-33;  
 Matches 172; Conservative 56; Mismatches 143; Indels 28; Gaps 12;

QY 2 VDEGALPEINSAARMYAGPGSASLVAAAOQWDSVASDLFSAASAFQSVWGLTVGSWIGS 61  
 Db 1 MDGVLPEINSGMYAGPGSAPMLAAAMDGATLQSTAAAYGVISVLT-GVWSGQ 59  
 QY 62 SAGLVAAASPYVAMSVTAGQAEITAAQVRVAAAAYETAVGLTPPPVIAENRAELMIL 121  
 Db 61 AASMAVTAAPITGMLYTBAKTIQTAIDAKAALAFEDQAVYMTLPPPVAAANKRIOLAL 120  
 QY 122 IATNLLGONTPAIVNAEYGEWMAODAAAFGAAATATATATLTFEEAPEMTSAGL 181  
 Db 121 IATNFGONTAIAATEAQYAEWMAODAAAFGAAATATATATLTFEEAPEMTSAGL 179  
 QY 182 LEQAAVEASDTPAAANQMLNNVPOALQLOAOT--QGTTPSSKLGIMKTYSPIRSP- 228  
 Db 180 TQAAAASQATDPLSL--LIEFTVQALQALITPSPIBEDFTFLDAFAGYATVGVQDVE 237  
 QY 239 SNVSM--ANNHSMNTSGVSMNTTSLSMKGPAPAAAQVQTAQNGVRASSISGSS 296  
 Db 238 SPVAGTIGAESNLGLINVDENPAEVTPEGFGELVSA--TSFGGVS-SCG- 289  
 QY 297 GSGGLGGVAAANLGRASVGSLSVPQMAAAAOAVTPAARLPLTSL--TSAERG-PGQ 353





**THIS PAGE BLANK (USPTO)**